

STIC-Biotech/ChemLib

154881

Mej

From: Dunston, Jennifer
Sent: Tuesday, May 31, 2005 10:53 AM
To: STIC-Biotech/ChemLib
Subject: Sequence Search 10/826523

Please do a sequence search for the nucleic acid sequences of SEQ ID NO: 40 and SEQ ID NO: 41 against the commercial and interference nucleotide databases.

The length of SEQ ID NO: 40 is 707 nucleotides. The length of SEQ ID NO: 41 is 3662 nucleotides. The sequences are related in that SEQ ID NO: 41 comprises the sequence of SEQ ID NO: 40.

Thank you.

Jennifer Dunston, Ph.D.
USPTO Art Unit 1636
REM 2B76
Mailbox: REM 2C70
(571) 272-2916

RECEIVED
MAY 31 2005
STIC

STAFF USE ONLY

Searcher: _____
Searcher Phone: 2-_____
Date Searcher Picked up: _____
Date Completed: 6-6
Searcher Prep/Rev. Time: _____
Online Time: _____

Type of Search

NA#: _____ AA#: _____
Interference: _____ SPDI: _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure#: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable

STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other(Specify): _____

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 5, 2005, 19:22:55 ; Search time 385.136 Seconds
(without alignments)
11284.911 Million cell updates/sec

Title: US-10-826-523-40
Perfect score: 707
Sequence: 1 ggaatccatcgctcaatttt.....aaaattgacgcatggatcc 707

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5706582 seqs, 3073711274 residues

Total number of hits satisfying chosen parameters: 11413164

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:
19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:
20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:
21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:
22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	707	100.0	707	13	US-10-001-189-40 Sequence 40, Appl
2	707	100.0	4613	13	US-10-001-189-46 Sequence 46, Appl
3	707	100.0	8999	13	US-10-001-189-48 Sequence 48, Appl
4	707	100.0	9012	13	US-10-001-189-49 Sequence 49, Appl
5	707	100.0	9013	13	US-10-001-189-50 Sequence 50, Appl
6	685	96.9	4943	13	US-10-001-189-54 Sequence 54, Appl
7	685	96.9	4944	13	US-10-001-189-55 Sequence 55, Appl
8	685	96.9	4944	13	US-10-001-189-56 Sequence 56, Appl
9	566.2	80.1	26565	16	US-10-085-959-91 Sequence 91, Appl
10	566.2	80.1	46819	14	US-10-114-170-72 Sequence 72, Appl
11	561.4	79.4	22306	14	US-10-114-170-251 Sequence 251, Appl

12	519.2	73.4	3662	13	US-10-001-189-41	Sequence 41, Appl
c 13	511	72.3	4941	13	US-10-001-189-53	Sequence 53, Appl
c 14	511	72.3	4951	13	US-10-001-189-51	Sequence 51, Appl
c 15	511	72.3	4952	13	US-10-001-189-52	Sequence 52, Appl
c 16	425.8	60.2	1427	18	US-10-363-345A-37517	Sequence 37517, A
c 17	425.8	60.2	1427	18	US-10-363-345A-37518	Sequence 37518, A
c 18	425.8	60.2	1427	18	US-10-363-345A-37517	Sequence 37517, A
c 19	425.8	60.2	1427	19	US-10-363-483A-37518	Sequence 37518, A
c 20	382.6	54.1	1427	18	US-10-363-345A-37519	Sequence 37519, A
c 21	382.6	54.1	1427	18	US-10-363-345A-37520	Sequence 37520, A
c 22	382.6	54.1	1427	19	US-10-363-483A-37519	Sequence 37519, A
c 23	382.6	54.1	1427	19	US-10-363-483A-37520	Sequence 37520, A
c 24	223	31.5	34063	14	US-10-114-170-96	Sequence 96, Appl
c 25	222.2	31.4	72480	17	US-10-418-837-2	Sequence 2, Appl
c 26	96	13.6	847	10	US-09-798-889-27	Sequence 27, Appl
c 27	96	13.6	847	17	US-10-633-680-27	Sequence 27, Appl
c 28	78	11.0	9984	13	US-10-001-189-63	Sequence 63, Appl
c 29	72.8	10.3	7670	13	US-10-001-189-57	Sequence 57, Appl
c 30	40	5.7	43	13	US-10-001-189-19	Sequence 19, Appl
c 31	40	5.7	3089	18	US-10-425-115-49726	Sequence 49726, A
c 32	39.8	5.6	676	18	US-10-425-115-91844	Sequence 91844, A
c 33	39.8	5.6	1673	17	US-10-425-114-21174	Sequence 21174, A
c 34	39.8	5.6	1796	17	US-10-425-114-29053	Sequence 29053, A
c 35	39.8	5.6	1872	18	US-10-425-115-91842	Sequence 91842, A
c 36	39	5.5	39	13	US-10-001-189-16	Sequence 16, Appl
c 37	38.4	5.4	1473	17	US-10-282-122A-20134	Sequence 20134, A
c 38	36.8	5.2	1487	18	US-10-437-963-91699	Sequence 91699, A
c 39	36.2	5.1	549	19	US-10-660-811A-23	Sequence 23, Appl
c 40	36.2	5.1	1403	18	US-10-767-701-13724	Sequence 13724, A
c 41	36.2	5.1	84830	19	US-10-660-811A-2	Sequence 2, Appl
c 42	36	5.0	1246	17	US-10-369-493-27114	Sequence 27114, A
c 43	35.4	5.0	2073	11	US-09-758-759-158	Sequence 158, App
c 44	35.4	5.0	96587	11	US-09-997-722-250	Sequence 250, App
c 45	35.4	5.0	109519	11	US-09-758-759-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1

US-10-001-189-40
; Sequence 40, Application US/10001189
; Publication No. US20020173634A1
; GENERAL INFORMATION:
; APPLICANT: FRASER JR., MALCOLM J.
; APPLICANT: LI, XU
; APPLICANT: BEAM, TERESA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
; FILE REFERENCE: VECTOR PIGGYBAC
; CURRENT APPLICATION NUMBER: US/10/001,189
; CURRENT FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/244,984
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/244,677
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 40
; LENGTH: 707
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: ITR Cartridge
; OTHER INFORMATION: sequence
US-10-001-189-40

Query Match 100.0%; Score 707; DB 13; Length 707;
Best Local Similarity 100.0%; Pred. No. 1.8e-235;
Matches 707; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATCCCATCGTCAATTTTACGACACTATCTTTCTAGGGTTAATCTAGCTCATCAGG 60

```
Db 1 G G A T C C C A T G C G T C A A T T T T A G C C A G A C T A T C T T T C T A G G T T A A T C T A G C T C A T C A G G 60
Qy 61 A T C A T A T C G T C G G G T C T T T T T C C G G C T C A G T C A T C G C C A A G C T G G C G T A T C T G G G C A 120
Db 61 A T C A T A T C G T C G G G T C T T T T T C C G G C T C A G T C A T C G C C A A G C T G G C G T A T C T G G G C A 120
Qy 121 T C G G G A G A A G A A G C C G T G C C T T T T C C G C G A G T T G A A G C G G C A T G A A A G A G T T T G 180
Db 121 T C G G G A G A A G A A G C C G T G C C T T T T C C G C G A G T T G A A G C G G C A T G A A A G A G T T T G 180
Qy 181 C C G A G G A T G A C T G C T G C T G C A T T G A C G T T T G A G C G A A A A C G C A G T T T A C C A T G A T T C 240
Db 181 C C G A G G A T G A C T G C T G C T G C A T T G A C G T T T G A G C G A A A A C G C A G T T T A C C A T G A T T C 240
Qy 241 G G A A A G T G T G C C A T G C A C G C C T T T A A C G T G A A C T G T T C G T T C A G G C C A C C T G G G A T A 300
Db 241 G G A A A G T G T G C C A T G C A C G C C T T T A A C G T G A A C T G T T C G T T C A G G C C A C C T G G G A T A 300
Qy 301 C C A G T T C G T C G C G G T T T T C C G G A C A C A G T T C C G G A T G T T C A G C C C G A A G C G C A T C A G C A 360
Db 301 C C A G T T C G T C G C G G T T T T C C G G A C A C A G T T C C G G A T G T T C A G C C C G A A G C G C A T C A G C A 360
Qy 361 A C C G A A C A A T A C C G G C G A C A G C C G G A A C T G C C G T G C C G T G T G C A G A T T A A T G A C A G C G 420
Db 361 A C C G A A C A A T A C C G G C G A C A G C C G G A A C T G C C G T G T G C A G A T T A A T G A C A G C G 420
Qy 421 G T C G C G C G T G G A T A T A C T C A G C G A G A C G G G T A T C T G C T G A T G C C C A G A A T 480
Db 421 G T C G C G C G T G G A T A T A C T C A G C G A G A C G G G T A T C T G C T G A T G C C C A G A A T 480
Qy 481 G G A C A T G A T A C C C G T G A G T T A C C G G C G G C G C C T C G T T C A T T C A G T T T T T G A A C 540
Db 481 G G A C A T G A T A C C C G T G A G T T A C C G G C G G C G C C T C G T T C A T T C A G T T T T T G A A C 540
Qy 541 C C G T G A G A C G G G C A G A C T C G C G T G C A A A T G T T T T A C A G C G T G A T G G A G C A G A T G A 600
Db 541 C C G T G A G A C G G G C A G A C T C G C G T G C A A A T G T T T T A C A G C G T G A T G G A G C A G A T G A 600
Qy 601 A G A T G C T C A C A C G C T G C A G A C A C G C A G C T A G A T T A A C C T A G A A G A T A A T C A T A T T G 660
Db 601 A G A T G C T C A C A C G C T G C A G A C A C G C A G C T A G A T T A A C C T A G A A G A T A A T C A T A T T G 660
Qy 661 T G A C G T A C G T T A A G A T A A T C A T G C G T A A A A T T G A C G C A T G G G A T C C 707
Db 661 T G A C G T A C G T T A A G A T A A T C A T G C G T A A A A T T G A C G C A T G G G A T C C 707
```

RESULT 2

US-10-001-189-46

; Sequence 46, Application US/10001189

; Publication No. US20020173634A1

; GENERAL INFORMATION:

; APPLICANT: FRASER JR., MALCOLM J.

; APPLICANT: LI, XU

; APPLICANT: BEAM, TERESA

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING

; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION

; TITLE OF INVENTION: VECTOR PIGGYBAC

; FILE REFERENCE: 835910-92098

; CURRENT APPLICATION NUMBER: US/10/001,189

; CURRENT FILING DATE: 2001-10-30

; PRIOR FILING DATE: 2000-11-01

; PRIOR FILING DATE: 2000-11-01

; PRIOR FILING DATE: 2000-10-31

; NUMBER OF SEQ ID NOS: 70

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 46

; LENGTH: 4613

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

```
; OTHER INFORMATION: Description of Artificial Sequence: pCRII-ITR
;
; OTHER INFORMATION: sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (344)..(922)
US-10-001-189-46
```

```
Query Match 100.0%; Score 707; DB 13; Length 4613;
Best Local Similarity 100.0%; Pred No. 4.9e-235;
Matches 707; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 G G A T C C C A T G C G T C A A T T T T A G C C A G A C T A T C T T T C T A G G T T A A T C T A G C T C A T C A G G 60
Db 294 G G A T C C C A T G C G T C A A T T T T A G C C A G A C T A T C T T T C T A G G T T A A T C T A G C T C A T C A G G 353
Qy 61 A T C A T A T C G T C G G G T C T T T T T C C G G C T C A G T C A T C G C C A A G C T G G C G T A T C T G G G C A 120
Db 354 A T C A T A T C G T C G G G T C T T T T T C C G G C T C A G T C A T C G C C A A G C T G G C G T A T C T G G G C A 413
Qy 121 T C G G G A G A A G A A G C C G T G C C T T T T C C G C G A G T T G A A G C G G C A T G A A A G A G T T T G 180
Db 414 T C G G G A G A A G A A G C C G T G C C T T T T C C G C G A G T T G A A G C G G C A T G A A A G A G T T T G 473
Qy 181 C C G A G A T G A C T G C T G C T G C A T T G A C G T T T G A G C G A A A A C G C A G T T T A C A T G A T T C 240
Db 474 C C G A G A T G A C T G C T G C T G C A T T G A C G T T T G A G C G A A A A C G C A G T T T A C A T G A T T C 533
Qy 241 G G G A A G G T G T G C C A T G C A C G C C T T T A A C G G T G A A C T G T T C G T T C A G G C C A C C T G G G A T A 300
Db 534 G G G A A G G T G T G C C A T G C A C G C C T T T A A C G G T G A A C T G T T C G T T C A G G C C A C C T G G G A T A 593
Qy 301 C C A G T T C G T C G C G G T T T T C C G A C A C A G T T C C G A T G T C A G C C C G A A G C G C A T C A G C A 360
Db 594 C C A G T T C G T C G C G G T T T T C C G A C A C A G T T C C G A T G T C A G C C C G A A G C G C A T C A G C A 653
Qy 361 A C C G A A C A A T A C C C G C G A C A G C C G G A A C T G C C G T G C C G T G T G C A G A T T A A T G A C A G C G 420
Db 654 A C C G A A C A A T A C C C G C G A C A G C C G G A A C T G C C G T G C C G T G T G C A G A T T A A T G A C A G C G 713
Qy 421 G T C G C G C G T G G A T A T A C G T C A G C G A G A C G G G T A T C T G C T G A T G C C C A G A A T 480
Db 714 G T C G C G C G T G G A T A T A C G T C A G C G A G A C G G G T A T C T G C T G A T G C C C A G A A T 773
Qy 481 G G A C A T G A T A C C C G T G A G T T A C C C G C G G C G C C C T C G T T C A T T C A G T T T T T G A A C 540
Db 774 G G A C A T G A T A C C C G T G A G T T A C C C G C G G C G C C C T C G T T C A T T C A G T T T T T G A A C 833
Qy 541 C C G T G A G A C C G G G C A G A C T C G C G G T G C A A A T G T G T T T T A C A G C G T G A T G G A G C A G A T G A 600
Db 834 C C G T G A G A C C G G G C A G A C T C G C G G T G C A A A T G T G T T T T A C A G C G T G A T G G A G C A G A T G A 893
Qy 601 A G A T G C T C A C A C G C T G C A G A C A C G C A G C T A G A T T A A C C T A G A A G A T A A T C A T A T T G 660
Db 894 A G A T G C T C A C A C G C T G C A G A C A C G C A G C T A G A T T A A C C T A G A A G A T A A T C A T A T T G 953
Qy 661 T G A C G T A C G T T A A G A T A A T C A T G C C T A A A A T T G A C G C A T G G G A T C C 707
Db 954 T G A C G T A C G T T A A G A T A A T C A T G C C T A A A A T T G A C G C A T G G G A T C C 1000
```

RESULT 3

US-10-001-189-48/c

; Sequence 48, Application US/10001189

; Publication No. US20020173634A1

; GENERAL INFORMATION:

; APPLICANT: FRASER JR., MALCOLM J.

; APPLICANT: LI, XU

; APPLICANT: BEAM, TERESA

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING

; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION

; TITLE OF INVENTION: VECTOR PIGGYBAC

; FILE REFERENCE: 835910-92098

; CURRENT APPLICATION NUMBER: US/10/001,189

; CURRENT FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/244,984
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/244,677
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 48
; LENGTH: 8999
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: p(PZ)-Bac-EYFP
; OTHER INFORMATION: sequence
US-10-001-189-48

Query Match 100.0%; Score 707; DB 13; Length 8999;
Best Local Similarity 100.0%; Pred. No. 6.9e-235;
Matches 707; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGATCCCATCGCTCAATTTTACGAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60
DB 8950 GGATCCCATCGCTCAATTTTACGAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 8891
QY 61 ATCATATCTCGGGTCTTTTTCGGCTCAGTCATCGCCCAAGCTGGCGTATCTGGGCA 120
DB 8890 ATCATATCTCGGGTCTTTTTCGGCTCAGTCATCGCCCAAGCTGGCGTATCTGGGCA 8831
QY 121 TCGGGAGGAGAAAGCCGCTCTTTTCCGCGAGGTTGAAGCGGCATGGAAGAGTTTG 180
DB 8830 TCGGGAGGAGAAAGCCGCTCTTTTCCGCGAGGTTGAAGCGGCATGGAAGAGTTTG 8771
QY 181 CCGAGGATGATCTGCTGCTGATTTGACGTTGAGCGAAACCGCATCTTACCATGATTC 240
DB 8770 CCGAGGATGATCTGCTGCTGATTTGAGCGAAACCGCATCTTACCATGATTC 8711
QY 241 GCGAAGTGTTGGCATGACGCTTTTAAACGGTGAATCTGTTCCGTCAGCCACCTGGGATA 300
DB 8710 GCGAAGTGTTGGCATGACGCTTTTAAACGGTGAATCTGTTCCGTCAGCCACCTGGGATA 8651
QY 301 CCAGTTCTGTCGGCTTTTCCGAGACACAGTTCCGATGGTTCAGCCGAGCGCATCAGCA 360
DB 8650 CCAGTTCTGTCGGCTTTTCCGAGACACAGTTCCGATGGTTCAGCCGAGCGCATCAGCA 8591
QY 361 ACCCGAACATACCGGCGAGCGGAACTCCGTCGCGGTGTCAGATTAATGACAGCG 420
DB 8590 ACCCGAACATACCGGCGAGCGGAACTCCGTCGCGGTGTCAGATTAATGACAGCG 8531
QY 421 GTGCGGCTGGGATATTACGTCAGCGAGGACGGGTATCTTGGCTGGATCCCGAGAAAT 480
DB 8530 GTGCGGCTGGGATATTACGTCAGCGAGGACGGGTATCTTGGCTGGATCCCGAGAAAT 8471
QY 481 GGCATGGATACCCCGTACGTTACCCGCGGCGGCTCGTTCAATTCACGTTTGAAC 540
DB 8470 GGCATGGATACCCCGTACGTTACCCGCGGCGGCTCGTTCAATTCACGTTTGAAC 8411
QY 541 CCGTGGAGGACGGGACAGCTCGCGGTGCAAAATGTGTTTACAGCGTGTGAGCAGATGA 600
DB 8410 CCGTGGAGGACGGGACAGCTCGCGGTGCAAAATGTGTTTACAGCGTGTGAGCAGATGA 8351
QY 601 AGATGTCGACGCTGCGAGACACGCGAGCTAGATTAAACCTAGAAAGATTAATCATTTG 660
DB 8350 AGATGTCGACGCTGCGAGACACGCGAGCTAGATTAAACCTAGAAAGATTAATCATTTG 8291
QY 661 TGACGTACGTTAAGATAATCATCGTAAATTTGACGCATGGATCC 707
DB 8290 TGACGTACGTTAAGATAATCATCGTAAATTTGACGCATGGATCC 8244

RESULT 4
US-10-001-189-49/c
; Sequence 49, Application US/10001189
; Publication No. US20020173634A1

; GENERAL INFORMATION:
; APPLICANT: FRASER JR., MALCOLM J.
; APPLICANT: LI, XU
; APPLICANT: BEAM, TERESA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
; TITLE OF INVENTION: VECTOR PIGGYBAC
; FILE REFERENCE: 835910-92098
; CURRENT APPLICATION NUMBER: US/10/001,189
; CURRENT FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/244,984
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/244,677
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 49
; LENGTH: 9012
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: p(PZ)-Bac-ECFP
; OTHER INFORMATION: sequence
US-10-001-189-49

Query Match 100.0%; Score 707; DB 13; Length 9012;
Best Local Similarity 100.0%; Pred. No. 6.9e-235;
Matches 707; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGATCCCATCGCTCAATTTTACGAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60
DB 8963 GGATCCCATCGCTCAATTTTACGAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 8904
QY 61 ATCATATCTCGGGTCTTTTTCGGCTCAGTCATCGCCCAAGCTGGCGTATCTGGGCA 120
DB 8903 ATCATATCTCGGGTCTTTTTCGGCTCAGTCATCGCCCAAGCTGGCGTATCTGGGCA 8844
QY 121 TCGGGAGGAGAAAGCCGCTCTTTTCCGCGAGGTTGAAGCGGCATGGAAGAGTTTG 180
DB 8843 TCGGGAGGAGAAAGCCGCTCTTTTCCGCGAGGTTGAAGCGGCATGGAAGAGTTTG 8784
QY 181 CCGAGGATGATCTGCTGCTGATTTGACGTTGAGCGAAACCGCATCTTACCATGATTC 240
DB 8783 CCGAGGATGATCTGCTGCTGATTTGAGCGAAACCGCATCTTACCATGATTC 8724
QY 241 GCGAAGTGTTGGCATGACGCTTTTAAACGGTGAATCTGTTCCGTCAGCCACCTGGGATA 300
DB 8723 GCGAAGTGTTGGCATGACGCTTTTAAACGGTGAATCTGTTCCGTCAGCCACCTGGGATA 8664
QY 301 CCAGTTCTGTCGGCTTTTCCGAGACACAGTTCCGGATGGTTCAGCCGAGCGCATCAGCA 360
DB 8663 CCAGTTCTGTCGGCTTTTCCGAGACACAGTTCCGGATGGTTCAGCCGAGCGCATCAGCA 8604
QY 361 ACCCGAACATACCGGCGAGCGGAACTCCGTCGCGGTGTCAGATTAATGACAGCG 420
DB 8603 ACCCGAACATACCGGCGAGCGGAACTCCGTCGCGGTGTCAGATTAATGACAGCG 8544
QY 421 GTGCGGCTGGGATATTACGTCAGCGAGGACGGGTATCTTGGCTGGATCCCGAGAAAT 480
DB 8543 GTGCGGCTGGGATATTACGTCAGCGAGGACGGGTATCTTGGCTGGATCCCGAGAAAT 8484
QY 481 GGCATGGATACCCCGTACGTTACCCGCGGCGGCTCGTTCAATTCACGTTTGAAC 540
DB 8483 GGCATGGATACCCCGTACGTTACCCGCGGCGGCTCGTTCAATTCACGTTTGAAC 8424
QY 541 CCGTGGAGGACGGGACAGCTCGCGGTGCAAAATGTGTTTACAGCGTGTGAGCAGATGA 600
DB 8423 CCGTGGAGGACGGGACAGCTCGCGGTGCAAAATGTGTTTACAGCGTGTGAGCAGATGA 8364
QY 601 AGATGTCGACGCTGCGAGACACGCGAGCTAGATTAAACCTAGAAAGATTAATCATTTG 660
DB 8363 AGATGTCGACGCTGCGAGACACGCGAGCTAGATTAAACCTAGAAAGATTAATCATTTG 8304

Db 1079 ACCGGAACAATACCGGCGACAGCGGAACCTGCCGTGCCGGTGTGCAGATTAATGACAGCG 1138
Qy 421 GTGCGGCGTGGGATATTACGTACAGCGAGGACGGGTATCTCGGTGGATCCGCGAGAAAT 480
Db 1139 GTGCGGCGTGGGATATTACGTACAGCGAGGACGGGTATCTCGGTGGATCCGCGAGAAAT 1198
Qy 481 GGACATGGATACCGCGTGGATTTACCGGCGGCGCGCTCGTTCAATTCAGCTTTTGAAC 540
Db 1199 GGACATGGATACCGCGTGGATTTACCGGCGG-----CTCGTTCAATTCAGCTTTTGAAC 1252
Qy 541 CCGTGAGGACGGGACGACCTCGGCTGCAATGTGTTTTACAGCGTGTGAGGACAGATGA 600
Db 1253 CCGTGAGGACGGGACGACCTCGGCTGCAATGTGTTTTACAGCGTGTGAGGACAGATGA 1312
Qy 601 AGATGCTCGACACGCTGCGAGAACGCGAGCTAGATTAAACCTAGAAAGATAATCATATTG 660
Db 1313 AGATGCTCGACACGCTGCGAGAACGCGAGCTAGATTAAACCTAGAAAGATAATCATATTG 1372
Qy 661 TGACGTACGTTAAAGATAATCATGCGTAAATTTGACCGCATGGATCC 707
Db 1373 TGACGTACGTTAAAGATAATCATGCGTAAATTTGACCGCATGGATCC 1419

RESULT 7

US-10-001-189-55
; Sequence 55, Application US/10001189
; Publication No. US20020173634A1
; GENERAL INFORMATION:
; APPLICANT: FRASER JR., MALCOLM J.
; APPLICANT: LI, XU
; APPLICANT: BEAM, TERESA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
; FILE REFERENCE: 835910-92098
; CURRENT APPLICATION NUMBER: US/10/001,189
; CURRENT FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/244,984
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/244,677
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 55
; TYPE: DNA
; LENGTH: 4944
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: PBS-ITR-EGFP
; OTHER INFORMATION: sequence
US-10-001-189-55

Query Match 96.9%; Score 685; DB 13; Length 4944;
Best Local Similarity 99.2%; Pred. No. 2.4e-227;
Matches 701; Conservative 0; Mismatches 0; Indels 6; Gaps 1;
Qy 1 GGATCCCATCGCTCAATTTTACGCGAGACTATCTTTAGGGTTAATCTAGCTGCATCAGG 60
Db 719 GGATCCCATCGCTCAATTTTACGCGAGACTATCTTTAGGGTTAATCTAGCTGCATCAGG 778
Qy 61 ATCATATCGTCGGGTCTTTTTCGGGCTCAGTATCGCCCAAGCTGGCGCTATCTGGGCA 120
Db 779 ATCATATCGTCGGGTCTTTTTCGGGCTCAGTATCGCCCAAGCTGGCGCTATCTGGGCA 838
Qy 121 TCGGGGAGGAGAACGCCGCTCTTTTCCGCGAGGTTGAAGCGGCATGGAAGAGTTTG 180
Db 839 TCGGGGAGGAGAACGCCGCTCTTTTCCGCGAGGTTGAAGCGGCATGGAAGAGTTTG 898
Qy 181 CCAGAGATGACTGCTGCTGATTGACGTTGAGCGGAAACGACGCTTTTACCATCATGATTC 240
Db 899 CCAGAGATGACTGCTGCTGATTGACGTTGAGCGGAAACGACGCTTTTACCATCATGATTC 958
Qy 241 GGGAAAGGTGTGGCCATGACACGCTTTTAAACGGTGAACCTGTTCTGTCAGGCCACCTGGGATA 300

Db 959 GGGAAAGGTGTGGCCATGACGCTTTTAAACGGTGAACCTGTTCTGTCAGGCCACCTGGGATA 1018
Qy 301 CCAAGTTTCGTCCGGCTTTTCCGAGACACAGTTTCCGGATGGTTCAGCCCAAGCGCATCAGCA 360
Db 1019 CCAAGTTTCGTCCGGCTTTTCCGAGACACAGTTTCCGGATGGTTCAGCCCAAGCGCATCAGCA 1078
Qy 361 ACCCGAACAATACCGGCGACAGCGGAACTGCGCGTCCGGTGTGCAGATTAATGACAGCG 420
Db 1079 ACCCGAACAATACCGGCGACAGCGGAACTGCGCGTCCGGTGTGCAGATTAATGACAGCG 1138
Qy 421 GTGCGGCGTGGGATATTACGTACAGCGAGGACGGGTATCTCGGTGGATCCGCGAGAAAT 480
Db 1139 GTGCGGCGTGGGATATTACGTACAGCGAGGACGGGTATCTCGGTGGATCCGCGAGAAAT 1198
Qy 481 GGACATGGATACCGCGTGGATTTACCGGCGGCGCGCTCGTTCAATTCAGCTTTTGAAC 540
Db 1199 GGACATGGATACCGCGTGGATTTACCGGCGG-----CTCGTTCAATTCAGCTTTTGAAC 1252
Qy 541 CCGTGAGGACGGGACGACCTCGGCTGCAATGTGTTTTACAGCGTGTGAGGACAGATGA 600
Db 1253 CCGTGAGGACGGGACGACCTCGGCTGCAATGTGTTTTACAGCGTGTGAGGACAGATGA 1312
Qy 601 AGATGCTCGACACGCTGCGAGAACGCGAGCTAGATTAAACCTAGAAAGATAATCATATTG 660
Db 1313 AGATGCTCGACACGCTGCGAGAACGCGAGCTAGATTAAACCTAGAAAGATAATCATATTG 1372
Qy 661 TGACGTACGTTAAAGATAATCATGCGTAAATTTGACCGCATGGATCC 707
Db 1373 TGACGTACGTTAAAGATAATCATGCGTAAATTTGACCGCATGGATCC 1419

RESULT 8

US-10-001-189-56
; Sequence 56, Application US/10001189
; Publication No. US20020173634A1
; GENERAL INFORMATION:
; APPLICANT: FRASER JR., MALCOLM J.
; APPLICANT: LI, XU
; APPLICANT: BEAM, TERESA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
; FILE REFERENCE: 835910-92098
; CURRENT APPLICATION NUMBER: US/10/001,189
; CURRENT FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/244,984
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/244,677
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 56
; LENGTH: 4944
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: PBS-ITR-EYFP
; OTHER INFORMATION: sequence
US-10-001-189-56

Query Match 96.9%; Score 685; DB 13; Length 4944;
Best Local Similarity 99.2%; Pred. No. 2.4e-227;
Matches 701; Conservative 0; Mismatches 0; Indels 6; Gaps 1;
Qy 1 GGATCCCATCGCTCAATTTTACGCGAGACTATCTTTAGGGTTAATCTAGCTGCATCAGG 60
Db 719 GGATCCCATCGCTCAATTTTACGCGAGACTATCTTTAGGGTTAATCTAGCTGCATCAGG 778
Qy 61 ATCATATCGTCGGGTCTTTTTCGGGCTCAGTATCGCCCAAGCTGGCGCTATCTGGGCA 120
Db 779 ATCATATCGTCGGGTCTTTTTCGGGCTCAGTATCGCCCAAGCTGGCGCTATCTGGGCA 838


```
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 960296.95017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 251-5000
; TELEFAX: (608) 251-9166
; INFORMATION FOR SEQ ID NO: 72:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 46819
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 72:
US-10-114-170-72

Query Match      80.1%; Score 566.2; DB 14; Length 46819;
Best Local Similarity 97.8%; Pred. No. 2.3e-185;
Matches 574; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGGCTCAGTCATCGCCCAAGC 104
DB 14795 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGGCTCAGTCATCGCCCAAGC 14854

QY 105 TGGCGCTATCTGGGCTCGGGAGGAGAACCCGTCCTTTTCCCGGAGGTTGAAGCG 164
DB 14855 TGGCGCTATCTGGGCTCGGGAGGAGAACCCGTCCTTTTCCCGGAGGTTGAAGCG 14914

QY 165 GCATGGAAGAGTTTGGCGGAGGATGACTGCTGCTGCAATTCAGCTTGAGCGAAGAACGCGC 224
DB 14915 GCATGGAAGAGTTTGGCGGAGGATGACTGCTGCTGCAATTCAGCTTGAGCGAAGAACGCGC 14974

QY 225 TTTACCATGATGATTCGGGAGGTTGGCCATGCACGCCCTTTTAAACCGTGAACCTGTTGCTT 284
DB 14975 TTTACCATGATGATTCGGGAGGTTGGCCATGCACGCCCTTTTAAACCGTGAACCTGTTGCTT 15034

QY 285 CAGGCCACCTGGGATACCACTGCTGCGCGCTTTTCGGGACACAGCTTCGGGATGGTCAGC 344
DB 15035 CAGGCCACCTGGGATACCACTGCTGCGCGCTTTTCGGGACACAGCTTCGGGATGGTCAGC 15094

QY 345 CCGAAGGCGATCAGCAACCGAACAATACCGGGACAGCGGAACTCCGCTGCGGTTG 404
DB 15095 CCGAAGGCGATCAGCAACCGAACAATACCGGGACAGCGGAACTCCGCTGCGGTTG 15154

QY 405 CAGATTAAATGACAGCGGTGGCGCTGGGATATTTACGTCAGCGAGGAGCGGTATCTCTGGC 464
DB 15155 CAGATTAAATGACAGCGGTGGCGCTGGGATATTTACGTCAGCGAGGAGCGGTATCTCTGGC 15214

QY 465 TGGATGCGGAGAAAATGGAATGGAATACCCCGTGAAGTTACCCGCGGCGCGCTGCTTTC 524
DB 15215 TGGATGCGGAGAAAATGGAATGGAATACCCCGTGAAGTTACCCGCGGCGCGCTGCTTTC 15274

QY 525 ATTACGCTTTTGAACCCGTGGAGGAGCGGACACTCGCGGTGCAATGTTTACAGC 584
DB 15275 ATTACGCTTTTGAACCCGTGGAGGAGCGGACACTCGCGGTGCAATGTTTACAGC 15334

QY 585 GTGATGAGCAGATGAAGATGCTCGACACGCTGCGAAGAACGCGAGCT 631
DB 15335 GTGATGAGCAGATGAAGATGCTCGACACGCTGCGAAGAACGCGAGCT 15381
```

RESULT 11

```
US-10-114-170-251/c
; Sequence 251, Application US/10114170
; Publication No. US20030023075A1
; GENERAL INFORMATION:
; APPLICANT: Blattner, Frederick R.
;            Burland, Valerie
;            Perna, Nicole T.
;            Plunkett, Guy
;            Welch, Rod
; TITLE OF INVENTION: No. US20030023075A1el Sequences of E. coli O157
; NUMBER OF SEQUENCES: 265
```

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch. 1.44Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 8.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/114.170
; FILING DATE: 01-Apr-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/453,702
; FILING DATE: 03-DEC-1999
; APPLICATION NUMBER: 60/110,955
; FILING DATE: 04-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 960296.95017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 251-5000
; TELEFAX: (608) 251-9166
; INFORMATION FOR SEQ ID NO: 251:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22306
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 251:
US-10-114-170-251

Query Match      79.4%; Score 561.4; DB 14; Length 22306;
Best Local Similarity 97.3%; Pred. No. 7.3e-184;
Matches 571; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGGCTCAGTCATCGCCCAAGC 104
DB 17193 ATCCAGTTCATCAGGATCATATCGTCGGGTCTTTTTCGGGCTCAGTCATCGCCCAAGC 17134

QY 105 TGGCGCTATCTGGGCTCGGGAGGAGAACCCGTCCTTTTCCCGGAGGTTGAAGCG 164
DB 17133 TGGCGCTATCTGGGCTCGGGAGGAGAACCCGTCCTTTTCCCGGAGGTTGAAGCG 17074

QY 165 GCATGGAAGAGTTTGGCGGAGGATGACTGCTGCTGCAATTCAGCTTGAGCGAAGAACGCGC 224
DB 17073 GCATGGAAGAGTTTGGCGGAGGATGACTGCTGCTGCAATTCAGCTTGAGCGAAGAACGCGC 17014

QY 225 TTTACCATGATGATTCGGGAGGTTGGCCATGCACGCCCTTTTAAACCGTGAACCTGTTGCTT 284
DB 17013 TTTACCATGATGATTCGGGAGGTTGGCCATGCACGCCCTTTTAAACCGTGAACCTGTTGCTT 16954

QY 285 CAGGCCACCTGGGATACCACTGCTGCGCGCTTTTCCGGACACAGCTTCGGGATGGTCAGC 344
DB 16953 CAGGCCACCTGGGATACCACTGCTGCGCGCTTTTCCGGACACAGCTTCGGGATGGTCAGC 16894

QY 345 CCGAAGGCGATCAGCAACCGAACAATACCGGGACAGCGGAACTCCCGTGGCGGTTG 404
DB 16893 CCGAAGGCGATCAGCAACCGAACAATACCGGGACAGCGGAACTCCCGTGGCGGTTG 16834

QY 405 CAGATTAAATGACAGCGGTGGCGCTGGGATATTTACGTCAGCGAGGAGCGGTATCTCTGGC 464
DB 16833 CAGATTAAATGACAGCGGTGGCGCTGGGATATTTACGTCAGCGAGGAGCGGTATCTCTGGC 16774

QY 465 TGGATGCGGAGAAAATGGAATGGAATACCCCGTGAAGTTACCCCGGCGCGCTGCTTTC 524
DB 16773 TGGATGCGGAGAAAATGGAATGGAATACCCCGTGAAGTTACCCCGGCGCGCTGCTTTC 16714
```

QY 525 ATTCAGCTTTTGAACCCGTGGAGGACGGCAGACTCGGGTGCATGTTGTTTACAGC 584
Db 16713 ATTCAGCTCTTTGAACCCGTGGAGGACGGCAGACCCGGGTGCAATGTTTACAGC 16654
QY 585 GTGATGAGCAGATGAAGATGCTCGACAGCTCGACAGCAACACGAGCT 631
Db 16653 GTGATGAGCAGATGAAGATGCTCGACAGCTCGACAGCAACACGAGCT 16607

RESULT 12

US-10-001-189-41
; Sequence 41, Application US/10001189
; Publication No. US20020173634A1
; GENERAL INFORMATION:
; APPLICANT: FRASER JR., MALCOLM J.
; APPLICANT: LI, XU
; APPLICANT: BEAM, TERESA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSDUCTION USING
; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
; TITLE OF INVENTION: VECTOR PIGGYBAC
; FILE REFERENCE: 835910-92098
; CURRENT APPLICATION NUMBER: US/10/001,189
; PRIOR FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/244,984
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/244,677
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 3662
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: pXL-Bac
; OTHER INFORMATION: sequence
US-10-001-189-41

Query Match 73.4%; Score 519.2; DB 13; Length 3662;
Best Local Similarity 98.5%; Pred. No. 1.5e-169;
Matches 524; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 2 GATCCCATGCTCAATTTTACGAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGGA 61
Db 983 GATCCCATGCTCAATTTTACGAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGGA 1042
QY 62 TCATATCGTCGGGTCTTTTTTCCGGCTCAGTCAATCGCCCAAGCTGGCGCTATCTGGGCAT 121
Db 1043 TCATATCGTCGGGTCTTTTTTCCGGCTCAGTCAATCGCCCAAGCTGGCGCTATCTGGGCAT 1102
QY 122 CGGGGAGGAAGACCGCGCTCTTTTCCCGGAGGTTGAAGCGGCATGGAAGAGTTTC 181
Db 1103 CGGGGAGGAAGACCGCGCTCTTTTCCCGGAGGTTGAAGCGGCATGGAAGAGTTTC 1162
QY 182 CGAGGATGACTGCTGCTGATTCACGTTGAGCGAAGACGACCTTTACCATGATGATTCG 241
Db 1163 CGAGGATGACTGCTGCTGATTCACGTTGAGCGAAGACGACCTTTACCATGATGATTCG 1222
QY 242 GGAAGGTGTGGCCATGCAACCGCTTTTAAACGGTGAACCTGTTTCAGGCCACCTGGGATAC 301
Db 1223 GGAAGGTGTGGCCATGCAACCGCTTTTAAACGGTGAACCTGTTTCAGGCCACCTGGGATAC 1282
QY 302 CAGTTGTCGGGCTTTCCGGACACAGTTCCGGATGGTCAGCCCGAAGCGATCAGCA 361
Db 1283 CAGTTGTCGGGCTTTCCGGACACAGTTCCGGATGGTCAGCCCGAAGCGATCAGCA 1342
QY 362 CCCGAAACATACCGGCGACAGCGGAACTCCGCTGCCGCTGTCAGATTAATGACACGCG 421
Db 1343 CCCGAAACATACCGGCGACAGCGGAACTCCGCTGCCGCTGTCAGATTAATGACACGCG 1402
QY 422 TCGCGGCTGGGATATTACGTCAGCGAGGACGGGTATCTTGGCTGGATGCCGAGAAATG 481

Db 1403 TCGCGGCTGGGATATTACGTCAGCGAGGACGGGTATCTTGGCTGGATGCCGAGAAATG 1462
QY 482 GACATGGATACCCGTCGAGTTACCCGGCGCGCGCTCGTTTCATTCACGTT 533
Db 1463 GACATGGATACCCGTCGAGTTACCCGGCGCGCGCTTGGCTGAATATCATGGT 1514

RESULT 13

US-10-001-189-53/c
; Sequence 53, Application US/10001189
; Publication No. US20020173634A1
; GENERAL INFORMATION:
; APPLICANT: FRASER JR., MALCOLM J.
; APPLICANT: LI, XU
; APPLICANT: BEAM, TERESA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSDUCTION USING
; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
; TITLE OF INVENTION: VECTOR PIGGYBAC
; FILE REFERENCE: 835910-92098
; CURRENT APPLICATION NUMBER: US/10/001,189
; PRIOR FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/244,984
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/244,677
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 53
; LENGTH: 4941
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: pXL-Bac-ECFP
; OTHER INFORMATION: sequence
US-10-001-189-53

Query Match 72.3%; Score 511; DB 13; Length 4941;
Best Local Similarity 100.0%; Pred. No. 1.3e-166;
Matches 511; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 CATGCTCAATTTTACGAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGGATCAT 66
Db 1129 CATGCTCAATTTTACGAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGGATCAT 1070
QY 67 TCGTGGGCTCTTTTTCGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGATCGGG 126
Db 1069 TCGTGGGCTCTTTTTCGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGATCGGG 1010
QY 127 AGGAAGAAGCCCGTGCCTTTTCCCGAGGTTGAAGCGGCATGGAAGAGTTTGGCGAGG 186
Db 1009 AGGAAGAAGCCCGTGCCTTTTCCCGAGGTTGAAGCGGCATGGAAGAGTTTGGCGAGG 950
QY 187 ATGACTGCTGTCATTTAGCTTTGAGCGAAGACGACCTTTACCATGATGATTCGGGAG 246
Db 949 ATGACTGCTGTCATTTAGCTTTGAGCGAAGACGACCTTTACCATGATGATTCGGGAG 890
QY 247 GTGTGCCATGCAACCGCTTTTAAACGGTGAACCTGTTTCAGGCCACCTGGGATACCATG 306
Db 889 GTGTGCCATGCAACCGCTTTTAAACGGTGAACCTGTTTCAGGCCACCTGGGATACCATG 830
QY 307 CGTGGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGCCCGAAGCGATCAGCAACCGGA 366
Db 829 CGTGGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGCCCGAAGCGATCAGCAACCGGA 770
QY 367 ACATATACCGCGACAGCCGGAACCTCCGCTGCCGCTGTCAGATTAATGACAGCGGTGCG 426
Db 769 ACAATATACCGCGACAGCCGGAACCTCCGCTGCCGCTGTCAGATTAATGACAGCGGTGCG 710
QY 427 CGCTGGATATTACGTCAGCGAGGACGGGTATCTTGGCTGGATGCCGAGAAATGACAT 486
Db 709 CGCTGGATATTACGTCAGCGAGGACGGGTATCTTGGCTGGATGCCGAGAAATGACAT 650
QY 487 GGATACCCCGTGAATTTACCCGCGCGCGCGC 517


```
Db 649 GGATACCCCGTGAGTTACCCGCGGCGCGC 619
|||||
RESULT 14
US-10-001-189-51/c
; Sequence 51, Application US/10001189
; Publication No. US20020173634A1
; GENERAL INFORMATION:
; APPLICANT: FRASER JR., MALCOLM J.
; APPLICANT: LI, XU
; APPLICANT: BEAM, TERESA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSDUCTION USING
; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
; FILE REFERENCE: 835910-92098
; CURRENT APPLICATION NUMBER: US/10/001,189
; PRIOR FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/244,984
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/244,677
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 51
; LENGTH: 4951
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: pXL-Bac-EYFP
US-10-001-189-51

Query Match 72.3%; Score 511; DB 13; Length 4951;
Best Local Similarity 100.0%; Pred. No. 1.3e-166;
Matches 511; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 CATGCGTCAATTTTACGACGACTATCTTTCTAGGGTTAACTAGTGCATCAGGATCATA 66
Db 1129 CATGCGTCAATTTTACGACGACTATCTTTCTAGGGTTAACTAGTGCATCAGGATCATA 1070

Qy 67 TCCTCGGGTCTTTTTCGGCTCAGTCATCGCCCAAGCTGGCGTATCTGGGATCGGG 126
Db 1069 TCCTCGGGTCTTTTTCGGCTCAGTCATCGCCCAAGCTGGCGTATCTGGGATCGGG 1010

Qy 127 AGAAGAAGCCGCTGCTTTTCCGCGAGGTGAAGCGCGCATGGAAGAGTTTGGCGAGG 186
Db 1009 AGAAGAAGCCGCTGCTTTTCCGCGAGGTGAAGCGCGCATGGAAGAGTTTGGCGAGG 950

Qy 187 ATGACTGCTGCTGCTGCTTTTAAAGCGAGTGAAGCGCGCATGGAAGAGTTTGGCGAGG 246
Db 949 ATGACTGCTGCTGCTGCTTTTAAAGCGAGTGAAGCGCGCATGGAAGAGTTTGGCGAGG 890

Qy 247 GTGTGGCCATGCAAGCTTTTAAAGCGAGTGAAGCGCGCATGGAAGAGTTTGGCGAGG 306
Db 889 GTGTGGCCATGCAAGCTTTTAAAGCGAGTGAAGCGCGCATGGAAGAGTTTGGCGAGG 830

Qy 307 CGTGGCGGCTTTTCCGCGACACAGTTCCGGATGTCAGCCGAGCGCATCAGCAACCCGA 366
Db 829 CGTGGCGGCTTTTCCGCGACACAGTTCCGGATGTCAGCCGAGCGCATCAGCAACCCGA 770

Qy 367 ACAATACCCGCGACAGCCGAACTGCGGTGTGCGAGTAAATGACAGCGGTGCGG 426
Db 769 ACAATACCCGCGACAGCCGAACTGCGGTGTGCGAGTAAATGACAGCGGTGCGG 710

Qy 427 CGCTGGGATTTACGTCAGCGAGGAGTATCCTGCTGGATGCGCGAGAAATGGACAT 486
Db 709 CGCTGGGATTTACGTCAGCGAGGAGTATCCTGCTGGATGCGCGAGAAATGGACAT 650

Qy 487 GGATACCCCGTGAGTTACCCGCGGCGCGC 517
Db 649 GGATACCCCGTGAGTTACCCGCGGCGCGC 619
```

Search completed: June 6, 2005, 04:27:41
Job time : 388.136 secs

```
RESULT 15
US-10-001-189-52/c
; Sequence 52, Application US/10001189
; Publication No. US20020173634A1
; GENERAL INFORMATION:
; APPLICANT: FRASER JR., MALCOLM J.
; APPLICANT: LI, XU
; APPLICANT: BEAM, TERESA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSDUCTION USING
; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
; FILE REFERENCE: 835910-92098
; CURRENT APPLICATION NUMBER: US/10/001,189
; PRIOR FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/244,984
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/244,677
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 52
; LENGTH: 4952
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: pXL-Bac-EGFP
US-10-001-189-52

Query Match 72.3%; Score 511; DB 13; Length 4952;
Best Local Similarity 100.0%; Pred. No. 1.3e-166;
Matches 511; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 CATGCGTCAATTTTACGACGACTATCTTTCTAGGGTTAACTAGTGCATCAGGATCATA 66
Db 1129 CATGCGTCAATTTTACGACGACTATCTTTCTAGGGTTAACTAGTGCATCAGGATCATA 1070

Qy 67 TCCTCGGGTCTTTTTCGGCTCAGTCATCGCCCAAGCTGGCGTATCTGGGATCGGG 126
Db 1069 TCCTCGGGTCTTTTTCGGCTCAGTCATCGCCCAAGCTGGCGTATCTGGGATCGGG 1010

Qy 127 AGAAGAAGCCGCTGCTTTTCCGCGAGGTGAAGCGCGCATGGAAGAGTTTGGCGAGG 186
Db 1009 AGAAGAAGCCGCTGCTTTTCCGCGAGGTGAAGCGCGCATGGAAGAGTTTGGCGAGG 950

Qy 187 ATGACTGCTGCTGCTGCTTTTAAAGCGAGTGAAGCGCGCATGGAAGAGTTTGGCGAGG 246
Db 949 ATGACTGCTGCTGCTGCTTTTAAAGCGAGTGAAGCGCGCATGGAAGAGTTTGGCGAGG 890

Qy 247 GTGTGGCCATGCAAGCTTTTAAAGCGAGTGAAGCGCGCATGGAAGAGTTTGGCGAGG 306
Db 889 GTGTGGCCATGCAAGCTTTTAAAGCGAGTGAAGCGCGCATGGAAGAGTTTGGCGAGG 830

Qy 307 CGTGGCGGCTTTTCCGCGACACAGTTCCGGATGTCAGCCGAGCGCATCAGCAACCCGA 366
Db 829 CGTGGCGGCTTTTCCGCGACACAGTTCCGGATGTCAGCCGAGCGCATCAGCAACCCGA 770

Qy 367 ACAATACCCGCGACAGCCGAACTGCGGTGTGCGAGTAAATGACAGCGGTGCGG 426
Db 769 ACAATACCCGCGACAGCCGAACTGCGGTGTGCGAGTAAATGACAGCGGTGCGG 710

Qy 427 CGCTGGGATTTACGTCAGCGAGGAGTATCCTGCTGGATGCGCGAGAAATGGACAT 486
Db 709 CGCTGGGATTTACGTCAGCGAGGAGTATCCTGCTGGATGCGCGAGAAATGGACAT 650

Qy 487 GGATACCCCGTGAGTTACCCGCGGCGCGC 517
Db 649 GGATACCCCGTGAGTTACCCGCGGCGCGC 619
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 5, 2005, 13:34:58 ; Search time 342.253 Seconds
(without alignments)
12228.534 Million cell updates/sec

Title: US-10-826-523-40
Perfect score: 707
Sequence: 1 ggatccatcgctcaatttt.....aaattgacgatggatcc 707

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870567 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: Geneseq_1980s:*
2: Geneseq_1990s:*
3: Geneseq_2000s:*
4: Geneseq_2001as:*
5: Geneseq_2001bs:*
6: Geneseq_2002as:*
7: Geneseq_2002bs:*
8: Geneseq_2003as:*
9: Geneseq_2003bs:*
10: Geneseq_2003cs:*
11: Geneseq_2003ds:*
12: Geneseq_2004as:*
13: Geneseq_2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	707	100.0	707	10	ACA55353 Eukaryoti
2	707	100.0	4613	10	ACA55360 transform
C 3	705.4	99.8	8999	10	ACA55361 transform
C 4	705.4	99.8	9009	10	ACA55362 transform
C 5	705.4	99.8	9012	10	ACA55363 transform
6	685	96.9	4943	10	ACA55369 transform
7	685	96.9	4943	10	ACA55367 transform
8	681.8	96.4	4944	10	ACA55368 transform
9	585.4	82.8	48502	12	ADP70043
C 10	580.6	82.1	4164	4	AAS44733
C 11	566.2	80.1	26565	6	ABH78924
C 12	566.2	80.1	26565	10	ADH80491
C 13	566.2	80.1	46819	9	ACD19059
14	564.6	79.9	46897	10	ADC00585
C 15	561.4	79.4	22306	9	ACD19238
C 16	561.4	79.4	91740	10	ADC00956
17	519.2	73.4	3661	10	ACA55354
C 18	511	72.3	4941	10	ACA55366
C 19	511	72.3	4952	10	ACA55365
C 20	499	70.6	4951	10	ACA55364

21	425.8	60.2	1427	6	ABQ50926
C 22	425.8	60.2	1427	6	ABQ50927
23	392.6	55.5	3822	5	AAS94485
C 24	382.6	54.1	1427	6	ABQ50928
25	382.6	54.1	1427	6	ABQ50929
C 26	223	31.5	34063	9	ACD19083
C 27	223	31.5	44029	10	ADC00710
28	223	31.5	49650	10	ADC00365
C 29	222.2	31.4	134141	6	ABN83487
C 30	171.2	24.2	822	5	AAS80174
C 31	96	13.6	847	2	AAS80174
C 32	78	11.0	9980	10	ACA55352
C 33	73.8	10.4	5194	4	AAC85595
C 34	73.8	10.4	5194	9	ADA09845
C 35	73.4	10.4	5194	4	AAC85596
C 36	73.4	10.4	5194	9	ADA09846
C 37	72.8	10.3	7670	10	ACA55351
C 38	67.8	9.6	2480	8	ABV76195
39	67.8	9.6	5679	4	AAC85602
40	67.8	9.6	5679	9	ADA09857
C 41	67.8	9.6	6723	4	AAC85593
C 42	67.8	9.6	6723	4	AAC85592
C 43	67.8	9.6	6723	9	ADA09843
C 44	67.8	9.6	6723	9	ADA09842
C 45	67.8	9.6	7560	4	AAC85599

ALIGNMENTS

RESULT 1
ACA55353
ID ACA55353 standard; DNA; 707 BP.
XX ACA55353;
XX
XX
XX 09-JUN-2003 (first entry)
XX
XX Eukaryotic transformation vector piggyBAC ITR cartridge.
XX
XX PiggyBac; transposon; eukaryotic transformation vector; ds;
XX transformed cell; transformed embryo; transgenic; ITR cartridge.
XX Synthetic.
XX
XX US2002173634-A1.
XX
XX 21-NOV-2002.
XX
XX 30-OCT-2001; 2001US-00001189.
XX
XX 31-OCT-2000; 2000US-0244677P.
XX
XX 01-NOV-2000; 2000US-0244984P.
XX
XX (FRAS/) FRASER M J.
XX (LIXX/) LI X.
XX (BEAM/) BEAM T.
XX (HUA/) HUA-VAN A.
XX
XX Fraser MJ, Li X, Beam T, Hua-Van A;
XX WPI; 2003-352597/33.
XX
XX New DNA molecule in the transposon piggyBac, useful for transferring
XX genes into host cells or embryos for transforming the cells of embryos
XX that can be used in making transgenic organisms.
XX
XX Example 3; Fig 3(C1); 151pp; English.
XX
XX The invention describes a DNA molecule comprising at least 163
XX consecutive nucleotide base pairs of the 3' terminal region beginning at
XX the 3' terminal base pair, and at least 125 consecutive nucleotide base
XX pairs of the 5' terminal region beginning at the 5' terminal base pair, of

CC the piggyBac molecule. The region extends from the restriction site SacI
CC to the end of the piggyBac molecule. The DNA molecule in the transposon
CC piggyBac is useful for transferring genes into host cells or embryos for
CC transforming the cells of embryos. The transposon cells or embryos are
CC useful for developing or making transgenic organisms. This sequence
CC represents a minimal sequence cartridge of the eukaryotic transformation
CC vector piggyBac
XX
SQ Sequence 707 BP; 164 A; 174 C; 208 G; 161 T; 0 U; 0 Other;

Query Match 100.0%; Score 707; DB 10; Length 707;
Best Local Similarity 100.0%; Pred. No. 2.6e-213;
Matches 707; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGATCCCATCGCTCAATTTTACGACGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60
DB 1 GGATCCCATCGCTCAATTTTACGACGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60
QY 61 ATCATATCGTCCGGTCTTTTTCGGGCTCAGTCAATGCCCAAGCTGGCGTATCTGGGCA 120
DB 61 ATCATATCGTCCGGTCTTTTTCGGGCTCAGTCAATGCCCAAGCTGGCGTATCTGGGCA 120
QY 121 TCGGGAGGAGGAAGCCGCTCTTTTCCGCGAGGTTGAAGCGGCATGGAAGAGTTTG 180
DB 121 TCGGGAGGAGGAAGCCGCTCTTTTCCGCGAGGTTGAAGCGGCATGGAAGAGTTTG 180
QY 181 CCGAGGATGACTGCTGCTGATTCAGCTTTCAGCGGAAACGCGTTTACCATGATGATTC 240
DB 181 CCGAGGATGACTGCTGCTGATTCAGCTTTCAGCGGAAACGCGTTTACCATGATGATTC 240
QY 241 GGGAGGATGCTGGCATCAGCCCTTTAAACGGTGAACCTGTTTCAGGCCACCTGGGATA 300
DB 241 GGGAGGATGCTGGCATCAGCCCTTTAAACGGTGAACCTGTTTCAGGCCACCTGGGATA 300
QY 301 CCAGTTCTCGCGCTTTTTCGGACACAGTTCCGGATGTCAGCCGAGCGATCAGCA 360
DB 301 CCAGTTCTCGCGCTTTTTCGGACACAGTTCCGGATGTCAGCCGAGCGATCAGCA 360
QY 361 ACCGACATACCGGAGCAGCGGAACTCCGTCGGCTGTCAGATTAATGACAGCG 420
DB 361 ACCGACATACCGGAGCAGCGGAACTCCGTCGGCTGTCAGATTAATGACAGCG 420
QY 421 GTGCGGCTGGGATATTCAGTTCAGCGAGGACGGTATCTTGGCTGGATCCCGCAGAAAT 480
DB 421 GTGCGGCTGGGATATTCAGTTCAGCGAGGACGGTATCTTGGCTGGATCCCGCAGAAAT 480
QY 481 GGACATGGATACCCCGTGAAGTTACCCGCGGCGCGCTCGTTCAATTCAGTTTGAAC 540
DB 481 GGACATGGATACCCCGTGAAGTTACCCGCGGCGCGCTCGTTCAATTCAGTTTGAAC 540
QY 541 CCGTGGAGGACGGCAGACTCGCGGTGCAATGTTTACAGCGTGATGGAGCAGATGA 600
DB 541 CCGTGGAGGACGGCAGACTCGCGGTGCAATGTTTACAGCGTGATGGAGCAGATGA 600
QY 601 AGATGCTCGACAGCTGCAGACACAGCAGCTAGATTAAACCTAGAAAGATAATCATATTG 660
DB 601 AGATGCTCGACAGCTGCAGACACAGCAGCTAGATTAAACCTAGAAAGATAATCATATTG 660
QY 661 TGACGTACGTTAAAGATAATCATGCGTAAATTTGACGATGGGATCC 707
DB 661 TGACGTACGTTAAAGATAATCATGCGTAAATTTGACGATGGGATCC 707

RESULT 2

ACA55360
ID ACA55360 standard; DNA; 4613 BP.

XX AC ACA55360;

XX AC ACA55360;

XX AC ACA55360;

XX AC ACA55360;

XX AC ACA55360;

XX AC ACA55360;

XX AC ACA55360;

XX AC ACA55360;

XX AC ACA55360;

XX AC ACA55360;

XX AC ACA55360;

KW PiggyBac; transposon; eukaryotic transformation vector; ds;
KW transformed cell; transposon; transgenic; plasmid; cyclic;
KW circular.
OS Synthetic.
XX US2002173634-A1.
XX 21-NOV-2002.

XX 30-OCT-2001; 2001US-00001189.

XX 31-OCT-2000; 2000US-0244677P.

XX 01-NOV-2000; 2000US-0244984P.

XX (FRAS/) FRASER M J.

XX (LIXX/) LI X.

XX (BEAM/) BEAM T.

XX (HUA/) HUA-VAN A.

XX Fraser MJ, Li X, Beam T, Hua-Van A;

XX MPI; 2003-352597/33.

XX P-PSDB; ABU70357, ABU70358.

XX New DNA molecule in the transposon piggyBac, useful for transferring

XX genes into host cells or embryos for transforming the cells of embryos

XX that can be used in making transgenic organisms.

XX Disclosure; Fig 10(B); 151pp; English.

XX The invention describes a DNA molecule comprising at least 163

XX consecutive nucleotide base pairs of the 3' terminal region beginning at

XX the 3' terminal base pair, and at least 125 consecutive nucleotide base

XX pairs of the 5' terminal region beginning at the 5' terminal base pair, of

XX the piggyBac molecule. The region extends from the restriction site SacI

XX to the end of the piggyBac molecule. The DNA molecule in the transposon

XX piggyBac is useful for transferring genes into host cells or embryos for

XX transforming the cells of embryos. The transposon cells or embryos are

XX useful for developing or making transgenic organisms. This sequence

XX represents plasmid used in the creation of minimal sequence eukaryotic

XX transformation vector piggyBac

XX Sequence 4613 BP; 1064 A; 1183 C; 1273 G; 1093 T; 0 U; 0 Other;

XX Query Match 100.0%; Score 707; DB 10; Length 4613;

XX Best Local Similarity 100.0%; Pred. No. 6.7e-213;

XX Matches 707; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATCCCATCGCTCAATTTTACGACGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60

DB 294 GGATCCCATCGCTCAATTTTACGACGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 353

QY 61 ATCATATCTCGGGTCTTTTTCGGGCTCAGTCAATGCCCAAGCTGGCGTATCTGGGCA 120

DB 354 ATCATATCTCGGGTCTTTTTCGGGCTCAGTCAATGCCCAAGCTGGCGTATCTGGGCA 413

QY 121 TCGGGAGGAGGAAGCCGCTCTTTTTCGGGAGGTTGAAGCGGCATGGAAGAGTTTG 180

DB 414 TCGGGAGGAGGAAGCCGCTCTTTTTCGGGAGGTTGAAGCGGCATGGAAGAGTTTG 473

QY 181 CCGAGGATGACTGCTGCTCATTTGAGCGGAAACGCGATTTACCATGATGATTC 240

DB 474 CCGAGGATGACTGCTGCTCATTTGAGCGGAAACGCGATTTACCATGATGATTC 533

QY 241 GGGAGGATGCTGGCATGCAAGCTTTTAAACGGTGAACCTGTTTACGGCCACCTGGGATA 300

DB 534 GGGAGGATGCTGGCATGCAAGCTTTTAAACGGTGAACCTGTTTACGGCCACCTGGGATA 593

QY 301 CCAGTTCTCGCGGCTTTTTCGGGACACAGTTCGGATGTCAGCCGAGGATGATTC 360

DB 594 CCAGTTCTCGCGGCTTTTTCGGGACACAGTTCGGATGTCAGCCGAGGATGATTC 653

Qy 361 ACCGGAACAATACCGCGACAGCGGAACTGCCGTCGGTGTGCAGATTAATGACAGCG 420
Db 654 ACCGGAACAATACCGCGACAGCGGAACTGCCGTCGGTGTGCAGATTAATGACAGCG 713
Qy 421 GTGCGCGCTGGGATATTACGTTCAGCGAGGACGCGGTATCTGCTGGATCCGCGAAT 480
Db 714 GTGCGCGCTGGGATATTACGTTCAGCGAGGACGCGGTATCTGCTGGATCCGCGAAT 773
Qy 481 GGACATGATACCCCGTGAAGTTACCCGCGGCGCGCTCGTTCAITTCAGTTTGAAC 540
Db 774 GGACATGATACCCCGTGAAGTTACCCGCGGCGCGCTCGTTCAITTCAGTTTGAAC 833
Qy 541 CCGTGAGGACGCGGACAGCTCGCGGTGCAAAATGTGTTTACAGGTGATGGAGCATGA 600
Db 834 CCGTGAGGACGCGGACAGCTCGCGGTGCAAAATGTGTTTACAGGTGATGGAGCATGA 893
Qy 601 AGATGCTCGACACGCTCGCAACACGCGAGCTAGATTAAACCTAGAAAGATAATCATATTG 660
Db 894 AGATGCTCGACACGCTCGCAACACGCGAGCTAGATTAAACCTAGAAAGATAATCATATTG 953
Qy 661 TGACGTACGTTAAAGATAATCATCGTAAATTTGACGATGGGATCC 707
Db 954 TGACGTACGTTAAAGATAATCATCGTAAATTTGACGATGGGATCC 1000

RESULT 3

ACA55361/c
ID ACA55361 standard; DNA; 8999 BP.

XX ACA55361;

DT 06-JUN-2003 (first entry)

XX Transformation vector piggyBAC related plasmid p(PZ)-Bac-EYFP.

XX PiggyBac; transposon; eukaryotic transformation vector; ds;

KW transformed cell; transformed embryo; transgenic; plasmid; cyclic;

KW circular.

XX Synthetic.

XX US2002173634-A1.

XX 21-NOV-2002.

XX 30-OCT-2001; 2001US-00001189.

XX 31-OCT-2000; 2000US-0244677P.

XX 01-NOV-2000; 2000US-0244984P.

XX (FRAS/) FRASER M J.

XX (LIXX/) LI X.

XX (BEAM/) BEAM T.

XX (HUAU/) HUA-VAN A.

XX Fraser MJ, Li X, Beam T, Hua-Van A;

XX WPI; 2003-352597/33.

XX Example 11; Fig 12(B); 151pp; English.

CC The invention describes a DNA molecule comprising at least 163
CC consecutive nucleotide base pairs of the 3' terminal region beginning at
CC the 3' terminal base pair, and at least 125 consecutive nucleotide base
CC pairs of the 5' terminal region beginning at the 5' terminal base pair, of
CC the piggyBac molecule. The region extends from the restriction site SacI
CC to the end of the piggyBac molecule. The DNA molecule in the transposon
CC .piggyBac is useful for transferring genes into host cells or embryos for
CC transforming the cells of embryos. The transformed cells or embryos are

CC useful for developing or making transgenic organisms. This sequence
CC represents plasmid used in the creation of minimal sequence eukaryotic
CC transformation vector piggyBac

SQ Sequence 8999 BP; 2180 A; 2416 C; 2342 G; 2061 T; 0 U; 0 Other;

Query Match 99.8%; Score 705.4; DB 10; Length 8999;

Best Local Similarity 99.9%; Pred. No. 3e-212;

Matches 706; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGATCCCATCGCTCAATTTTACGCACACTATCTTTCTAGGGTTAACTAGCTGCATCAGG 60

Db 8950 GGATCCCATCGCTCAATTTTACGCACACTATCTTTCTAGGGTTAACTAGCTGCATCAGG 8891

Qy 61 ATCATATCTGTCGGGTCTTTTTCGGCTCAGTCAATCGCCCAAGCTGCGCTATCTGGCA 120

Db 8890 ATCATATCTGTCGGGTCTTTTTCGGCTCAGTCAATCGCCCAAGCTGCGCTATCTGGCA 8831

Qy 121 TCGGGAGGAAGAAGCCCGTCTTTTCCGCGAGGTGAGCGGCATGGAAGAGTTTG 180

Db 8830 TCGGGAGGAAGAAGCCCGTCTTTTCCGCGAGGTGAGCGGCATGGAAGAGTTTG 8771

Qy 181 CCGAGGATGACTGCTGCTCATTTGAGCGGAAACGACGTTTACCATCATGATTC 240

Db 8770 CCGAGGATGACTGCTGCTCATTTGAGCGGAAACGACGTTTACCATCATGATTC 8711

Qy 241 GGGAGGTGTGGCCATGCACGCTTTAAACGCTGAACCTGTTCTGTTACAGCCACCTGGGATA 300

Db 8710 GGGAGGTGTGGCCATGCACGCTTTAAACGCTGAACCTGTTCTGTTACAGCCACCTGGGATA 8651

Qy 301 CCAGTTCTGCGGGCTTTTCCGGACACAGTTCGGATGTTACGCCGAGCGCATCAGCA 360

Db 8650 CCAGTTCTGCGGGCTTTTCCGGACACAGTTCGGATGTTACGCCGAGCGCATCAGCA 8591

Qy 361 ACCCGAACATACCCGCGACGACCGGAACTGCGTCCGCGTGTGCAGATTAAATGACAGCG 420

Db 8590 ACCCGAACATACCCGCGACGACCGGAACTGCGTCCGCGTGTGCAGATTAAATGACAGCG 8531

Qy 421 GTGCGCGCTGGGATATTACGTTCAGCGAGGACGGGTATCTGCTGGATGCCCGAGAAAT 480

Db 8530 GTGCGCGCTGGGATATTACGTTCAGCGAGGACGGGTATCTGCTGGATGCCCGAGAAAT 8471

Qy 481 GGACATGATACCCCGTGAAGTTACCCGCGGCGCGCTCGTTCAITTCAGTTTGAAC 540

Db 8470 GGACATGATACCCCGTGAAGTTACCCGCGGCGCGCTCGTTCAITTCAGTTTGAAC 8411

Qy 541 CCGTGAGGACGCGGACAGCTCCGCGTGCAAAATGTGTTTACAGCGGTGATGGAGCATGA 600

Db 8410 CCGTGAGGACGCGGACAGCTCCGCGTGCAAAATGTGTTTACAGCGGTGATGGAGCATGA 8351

Qy 601 AGATGCTCGACACGCTGCAGAACACGACGCTAGATTAAACCTAGAAAGATAATCATATTG 660

Db 8350 AGATGCTCGACACGCTGCAGAACACGACGCTAGATTAAACCTAGAAAGATAATCATATTG 8291

Qy 661 TGACGTACGTTAAAGATAATCATGCTAAATTTGACGATGGGATCC 707

Db 8290 TGACGTACGTTAAAGATAATCATGCTAAATTTGACGATGGGATCC 8244

RESULT 4

ACA55362/c

ID ACA55362 standard; DNA; 9009 BP.

XX ACA55362;

XX 06-JUN-2003 (first entry)

XX Transformation vector piggyBAC related plasmid p(PZ)-Bac-ECFP.

XX PiggyBac; transposon; eukaryotic transformation vector; ds;

KW transformed cell; transformed embryo; transgenic; plasmid; cyclic;

XX circular.

Query Match	99.8%	Score	705.4	DB	10	Length	9012
Best Local Similarity	99.9%	Pred. No.	3e-212				
Matches	706	Conservative	0	Mismatches	1	Indels	0
Qy	1	GGATCCCATCGTCGTCAAATTTTACGCAGACTATCTTTCTAGGGTTAAATCTAGCTGCATCAGG	60				
Db							
	8963	GGATCCCATCGTCGTCAAATTTTACGCAGACTATCTTTCTAGGGTTAAATCTAGCTGCATCAGG	8904				
Qy	61	ATCATATCGTCGGGCTCTTTTTCGGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA	120				
Db							
	8903	ATCATATCGTCGGGCTCTTTTTCGGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA	8844				
Qy	121	TCGGGGAGGAAGACCCGTCCTTTTCCCGCGAGGTTGAAGCGGCAATGGAAGAGGTTTG	180				
Db							
	8943	TCGGGGAGGAAGACCCGTCCTTTTCCCGCGAGGTTGAAGCGGCAATGGAAGAGGTTTG	8784				
Qy	181	CCGAGGATGACTGCTGCTGCATTGACGTTTCAGCGGAAACGCACGTTTACCATGATGATTC	240				
Db							
	8783	CCGAGGATGACTGCTGCTGCATTGACGTTTCAGCGGAAACGCACGTTTACCATGATGATTC	8724				
Qy	241	GGGAAGGTGGCCCATCAGCCCTTTTAAACGGTGAACTGTTCTGTTACGGCCACCTCGGGATA	300				
Db							
	8723	GGGAAGGTGGCCCATCAGCCCTTTTAAACGGTGAACTGTTCTGTTACGGCCACCTCGGGATA	8664				
Qy	301	CCAGTTCGTCGGGCTTTTTCGGACACAGTTTCGGATGGTCAGCCCGAAGGCATCAGCA	360				
Db							
	8663	CCAGTTCGTCGGGCTTTTTCGGACACAGTTTCGGATGGTCAGCCCGAAGGCATCAGCA	8604				
Qy	361	ACCGAACAAATACGGCGACAGCCGGAATCGCGTGCCGGTGTGCAGATTAATCACACGC	420				
Db							
	8603	ACCGAACAAATACGGCGACAGCCGGAATCGCGTGCCGGTGTGCAGATTAATCACACGC	8544				
Qy	421	GTGCGGCGCTGGGATATTAAGTCAGCAGGACCGGGTATCCTGGCTGGATGCCCGCAGAAAT	480				
Db							
	8543	GTGCGGCGCTGGGATATTAAGTCAGCAGGACCGGGTATCCTGGCTGGATGCCCGCAGAAAT	8484				
Qy	481	GGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCTCGTTCAATTCACGTTTTTGAAC	540				
Db							
	8483	GGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCTCGTTCAATTCACGTTTTTGAAC	8424				
Qy	541	CCGTGGAGGACGGGCAGACTCGCGGTCAAATGTGTTTTACGGGTGATGGAGCAGATGA	600				
Db							
	8423	CCGTGGAGGACGGGCAGACTCGCGGTCAAATGTGTTTTACGGGTGATGGAGCAGATGA	8364				
Qy	601	AGATGCTCGACACGCTCGAGAACACGCAGCTAGATTAAACCTAGAAAGATAATCATATTG	660				
Db							
	8363	AGATGCTCGACACGCTCGAGAACACGCAGCTAGATTAAACCTAGAAAGATAATCATATTG	8304				
Qy	661	TGACGTAAGTAAAGATAATCATGCGGTAAATTTGACGCATGGGATCC	707				
Db							
	8303	TGACGTAAGTAAAGATAATCATGCGGTAAATTTGACGCATGGGATCC	8257				

RESULT 6	
ACA55369	
ID	ACA55369 standard; DNA; 4943 BP.
XX	
XX	ACA55369;
XX	
DT	06-JUN-2003 (first entry)
XX	
DE	Transformation vector pIGS-BAC related plasmid pBS-TTR-EYFP.
XX	
XX	PiggyBac; transposon; eukaryotic transformation vector; ds;
KW	transformed cell; transgenic; plasmid; cyclic;
KW	circular.
XX	
OS	Synthetic.
XX	
PN	US2002173634-A1.
XX	
PD	21-NOV-2002.

XX		30-OCT-2001; 2001US-00001189.	
Pf		31-OCT-2000; 2000US-0244677P.	
XX		01-NOV-2000; 2000US-0244984P.	
PR		(FRAS/) FRASER M J.	
XX		(LIXX/) LI X.	
PA		(BEAM/) BEAM T.	
PA		(HUAV/) HUA-VAN A.	
XX			
PI		Fraser MJ, Li X, Beam T, Hua-Van A;	
XX		WPI; 2003-352597/33..	
DR			
XX		New DNA molecule in the transposon piggyBac, useful for transferring	
PT		genes into host cells or embryos for transforming the cells of embryos	
PT		that can be used in making transgenic organisms.	
XX			
FS		Example 6; Fig 20(B); 151pp; English.	
XX			
CC		The invention describes a DNA molecule comprising at least 163	
CC		consecutive nucleotide base pairs of the 3' terminal region beginning at	
CC		the 3' terminal base pair, and at least 125 consecutive nucleotide base	
CC		pairs of the 5' terminal region beginning at the 5'terminal base pair, of	
CC		the piggyBac molecule. The region extends from the restriction site SacI	
CC		to the end of the piggyBac molecule. The DNA molecule in the transposon	
CC		piggyBac is useful for transferring genes into host cells or embryos for	
CC		transforming the cells of embryos. The transformed cells or embryos are	
CC		useful for developing or making transgenic organisms. This sequence	
CC		represents plasmid used in the creation of minimal sequence eukaryotic	
CC		transformation vector piggyBac	
XX			
SQ		Sequence 4943 BP; 1220 A; 1302 C; 1245 G; 1176 T; 0 U; 0 Other;	
		Query March 96.9%; Score 685; DB 10; Length 4943;	
		Best Local Similarity 99.2%; Pred. No. 6.6e-206;	
		Matches 701; Conservative 0; Mismatches 0; Indels 6; Gaps 1;	
Oy	1	GGATCCCATCGGTCAATTTTTACGCAGACTATCTTTCTAGGGTTAACTAGCTGCATCAGG 60	
Db	719	GGATCCCATCGGTCAATTTTTACGCAGACTATCTTTCTAGGGTTAACTAGCTGCATCAGG 778	
Oy	61	ATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCTATCGCCCAAGCTGGCGTATCTGGCA 120	
Db	779	ATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCTATCGCCCAAGCTGGCGTATCTGGCA 838	
Oy	121	TGGGGAGGAAGAAGCCCGTGCTTTTTTCCGGCGAGGTTGAAGCGGANTGAAAAGAGTTTG 180	
Db	839	TGGGGAGGAAGAAGCCCGTGCTTTTTTCCGGCGAGGTTGAAGCGGANTGAAAAGAGTTTG 898	
Oy	181	CCGAGGATGACTGCTGCTGCAATTGAGTTTGAGCGGAAAAAAGCGGATGCAATGATTC 240	
Db	899	CCGAGGATGACTGCTGCTGCAATTGAGTTTGAGCGGAAAAAAGCGGATGCAATGATTC 958	
Oy	241	GGGAAAGTGTGGCCATGCAGCCCTTTAAACGGTGAACTGTTTCGTTCAGGCCACCCTGGATA 300	
Db	959	GGGAAGTGTGGCCATGCAGCCCTTTAAACGGTGAACTGTTTCGTTCAGGCCACCCTGGATA 1018	
Oy	301	CCAGTTCGTTCGGGCTTTTCCGGACACAGTTCCTCGGATGFTFACGCCGACCGCATAGCA 360	
Db	1019	CCAGTTCGTTCGGGCTTTTCCGGACACAGTTCCTCGGATGFTFACGCCGACCGCATAGCA 1078	
Oy	361	ACCCGACAAATACCGGCAGCCGGAACCTGCCGTGCCGTGTGCAGATTAATGACAGCG 420	
Db	1079	ACCCGAAACAATACCGGCAGCCGGAACCTGCCGTGCCGTGTGCAGATTAATGACAGCG 1138	
Oy	421	GTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGCTGGATGCCCGACAAAT 480	
Db	1139	GTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGCTGGATGCCCGACAAAT 1198	
Oy	481	GGACATGGATACCCCCTGAGTTTACCCGGCGGGCGCCCTCGTTCATTCAACGTTTTTGAAC 540	

Db 1199 GGACATGGATACCCCGTGAAGTACCCGGCGG-----CTGGTTCAATTCACGTTTGAAC 1252
Qy 541 CCCTGGAGGACGGCGAGACTCGCGGTGCAAAATGTTTACAGCGTGATGGACAGATGA 600
Db 1253 CCCTGGAGGACGGCGAGACTCGCGGTGCAAAATGTTTACAGCGTGATGGAGCAGATGA 1312
Qy 601 AGATGCTCGACACGCTGCGAGAACACGCGAGCTAGATTAAACCTTGAAGAATAATCATATTG 660
Db 1313 AGATGCTCGACACGCTGCGAGAACACGCGAGCTAGATTAAACCTTGAAGAATAATCATATTG 1372
Qy 661 TGACGTACGTTAAAGATAATCATCGCTAAATTTGACGCGATGGGATCC 707
Db 1373 TGACGTACGTTAAAGATAATCATCGCTAAATTTGACGCGATGGGATCC 1419

RESULT 7

ACA55367
ID ACA55367 standard; DNA; 4943 BP.

XX ACA55367;

DT 06-JUN-2003 (first entry)

XX Transformation vector piggyBAC related plasmid pBS-ITR-EGFP.

DE PiggyBac; transposon; eukaryotic transformation vector; ds;
KW transformed cell; transformed embryo; transgenic; plasmid; cyclic;
KW circular.

XX Synthetic.

XX US2002173634-A1.

XX 21-NOV-2002.

XX 30-OCT-2001; 2001US-00001189.

XX 31-OCT-2000; 2000US-0244677P.

XX 01-NOV-2000; 2000US-0244984P.

XX (FRAS/) FRASER M J.

XX (LIXX/) LI X.

XX (BEAM/) BEAM T.

XX (HUA/) HUA-VAN A.

XX Fraser MJ, Li X, Beam T, Hua-Van A;

XX WPI; 2003-352597/33.

XX New DNA molecule in the transposon piggyBac, useful for transferring
PT genes into host cells or embryos for transforming the cells of embryos
PT that can be used in making transgenic organisms.

XX Example 6; Fig 18(B); 151pp; English.

XX The invention describes a DNA molecule comprising at least 163
CC consecutive nucleotide base pairs of the 3' terminal region beginning at
CC the 3' terminal base pair, and at least 125 consecutive nucleotide base
CC pairs of the 5' terminal region beginning at the 5' terminal base pair, of
CC the piggyBac molecule. The region extends from the restriction site SacI
CC to the end of the piggyBac molecule. The DNA molecule in the transposon
CC piggyBac is useful for transforming genes into host cells or embryos for
CC transforming the cells of embryos. The transformed cells or embryos are
CC useful for developing or making transgenic organisms. This sequence
CC represents plasmid used in the creation of minimal sequence eukaryotic
CC transformation vector piggyBac

SQ Sequence 4943 BP; 1221 A; 1305 C; 1245 G; 1172 T; 0 U; 0 Other;

Query Match

Best Local Similarity 96.9%; Score 685; DB 10; Length 4943;

Matches 701; Conservative 0; Mismatches 0; Indels 6; Gaps 1;

Qy 1 GGATCCATCGCTCAATTTTACGACAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60
Db 719 GGATCCATCGCTCAATTTTACGACAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 778
Qy 61 ATCATATCGTCGGGTCTTTTTCGGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA 120
Db 779 ATCATATCGTCGGGTCTTTTTCGGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA 838
Qy 121 TCGGGGAGGAAGAACCGCTGCTTTTCCGCGAGGTTGAAGCGGCATCGAAGAGCTTTG 180
Db 839 TCGGGGAGGAAGAACCGCTGCTTTTCCGCGAGGTTGAAGCGGCATCGAAGAGCTTTG 898
Qy 181 CCGAGGATGACTGCTGCTGCTCAATTTGAGCGGAAAACGACGTTTACCATGATGATTC 240
Db 899 CCGAGGATGACTGCTGCTGCTCAATTTGAGCGGAAAACGACGTTTACCATGATGATTC 958
Qy 241 GGGAAAGGTGTGCCATGCAAGCTTTTAAACGGTGAATGTTCTTACGGCCACTCTGGGATA 300
Db 959 GGGAAAGGTGTGCCATGCAAGCTTTTAAACGGTGAATGTTCTTACGGCCACTCTGGGATA 1018
Qy 301 CCAGTTCGTGCGGGCTTTTCCGGACACAGTTCGGGATGGTCCAGCCGAGCGCATCAGCA 360
Db 1019 CCAGTTCGTGCGGGCTTTTCCGGACACAGTTCGGGATGGTCCAGCCGAGCGCATCAGCA 1078
Qy 361 ACCCGAAACAATACCGCGCAGACAGCCGGAACCTGCGCGGTGTGCAGATTAATGACAGCG 420
Db 1079 ACCCGAAACAATACCGCGCAGACAGCCGGAACCTGCGCGGTGTGCAGATTAATGACAGCG 1138
Qy 421 GTGCGGCGCTGGGATATTACGTCAGGAGGACGGGTATCTCTGGCTGGATGCCGAGAAAT 480
Db 1139 GTGCGGCGCTGGGATATTACGTCAGGAGGACGGGTATCTCTGGCTGGATGCCGAGAAAT 1198
Qy 481 GGACATGGATACCCCGTGAGTTACCCGGCGGCGCGCTCGTTTCATTACGTTTTTTGAAC 540
Db 1199 GGACATGGATACCCCGTGAGTTACCCGGCGGCGCGCTCGTTTCATTACGTTTTTTGAAC 1252
Qy 541 CCGTGAGGACCGGCGAGACTCGCGGTGCAAAATGTGTTTTACAGCGTGATGGAGCAGATGA 600
Db 1253 CCGTGAGGACCGGCGAGACTCGCGGTGCAAAATGTGTTTTACAGCGTGATGGAGCAGATGA 1312
Qy 601 AGATGCTCGACACGCTGCGAGAACACGCGAGCTAGATTAAACCTAGAAAGATAATCATATTG 660
Db 1313 AGATGCTCGACACGCTGCGAGAACACGCGAGCTAGATTAAACCTAGAAAGATAATCATATTG 1372
Qy 661 TGACGTACGTTAAAGATAATCATGCTAAATTTGACGCGATGGGATCC 707
Db 1373 TGACGTACGTTAAAGATAATCATGCTAAATTTGACGCGATGGGATCC 1419

RESULT 8

ACA55368
ID ACA55368 standard; DNA; 4944 BP.

XX ACA55368;

XX 06-JUN-2003 (first entry)

XX Transformation vector piggyBAC related plasmid pBS-ITR-EGFP.

XX PiggyBac; transposon; eukaryotic transformation vector; ds;
KW transformed cell; transformed embryo; transgenic; plasmid; cyclic;
KW circular.

XX Synthetic.

XX US2002173634-A1.

XX 21-NOV-2002.

XX 30-OCT-2001; 2001US-00001189.

XX 31-OCT-2000; 2000US-0244677P.

PR 01-NOV-2000; 2000US-0244984P.

XX (FRASER) FRASER M J.
PA (LIXX/) LI X.
PA (BEAM/) BEAM T.
PA (HUAU/) HUA-VAN A.
XX Fraser MJ, Li X, Beam T, Hua-Van A;
XX WPI; 2003-352597/33.
XX
XX New DNA molecule in the transposon piggyBac, useful for transferring
PT genes into host cells or embryos for transforming the cells of embryos
PT that can be used in making transgenic organisms.
XX
XX Example 6; Fig 19(B); 151pp; English.
XX
XX The invention describes a DNA molecule comprising at least 163
CC consecutive nucleotide base pairs of the 3' terminal region beginning at
CC the 3' terminal base pair, and at least 125 consecutive nucleotide base
CC pairs of the 5' terminal region beginning at the 5' terminal base pair, of
CC the piggyBac molecule. The region extends from the restriction site SacI
CC to the end of the piggyBac molecule. The DNA molecule in the transposon
CC piggyBac is useful for transferring genes into host cells or embryos for
CC transforming the cells of embryos. The transformed cells or embryos are
CC useful for developing or making transgenic organisms. This sequence
CC represents plasmid used in the creation of minimal sequence eukaryotic
CC transformation vector piggyBac
XX
XX Sequence 4944 BP; 1221 A; 1302 C; 1247 G; 1174 T; 0 U; 0 Other;
SQ
Query Match 96.4%; Score 681.8; DB 10; Length 4944;
Best Local Similarity 98.9%; Pred. No. 6.8e-205;
Matches 699; Conservative 0; Mismatches 2; Indels 6; Gaps 1;
QY 1 GGATCCCATCGTCGCAATTTTACGACAGACTATCTTTCTAGGGTTAACTAGTCGCATCAGG 60
DB 719 GGATCCCATCGTCGCAATTTTACGACAGACTATCTTTCTAGGGTTAACTAGTCGCATCAGG 778
QY 61 ATCATATCGTCGGGTCTTTTTCGGGTCTAGTCATCGCCCAAGTCGGCGTATCTGGGCA 120
DB 779 ATCATATCGTCGGGTCTTTTTCGGGTCTAGTCATCGCCCAAGTCGGCGTATCTGGGCA 838
QY 121 TCGGGAGGAGAGCCGCTGCTTTTCCGCGAGGTGAAGCGGCATGGAAGAGTTTG 180
DB 839 TCGGGAGGAGAGCCGCTGCTTTTCCGCGAGGTGAAGCGGCATGGAAGAGTTTG 898
QY 181 CCAGAGGATGACTGCTGCTGATTGACGTTTCAGCGGAAACGACGTTTACCATCATGATTC 240
DB 899 CCAGAGGATGACTGCTGCTGATTGACGTTTCAGCGGAAACGACGTTTACCATCATGATTC 958
QY 241 GGGAAAGGTGTGGCCATGCACGCTTTTAAACGGTGAACCTGTTTCAGGCCACCTGGGATA 300
DB 959 GGGAAAGGTGTGGCCATGCACGCTTTTAAACGGTGAACCTGTTTCAGGCCACCTGGGATA 1018
QY 301 CCAGTTGCTGCGGGCTTTTCCGACACAGTTCCGATGTCAGCCGAGCGGATCAGCA 360
DB 1019 CCAGTTGCTGCGGGCTTTTCCGACACAGTTCCGATGTCAGCCGAGCGGATCAGCA 1078
QY 361 ACCCGACATACCGGCGACGCGGAACTGCGGTGCGGTGTCAGATTAATGACAGCG 420
DB 1079 ACCCGACATACCGGCGACGCGGAACTGCGGTGCGGTGTCAGATTAATGACAGCG 1138
QY 421 GTGCGGCGCTGGGATATTAGCTCAGCGAGGACGGGTATCTGCTGGATGCCCGCAGAAAT 480
DB 1139 GTGCGGCGCTGGGATATTAGCTCAGCGAGGACGGGTATCTGCTGGATGCCCGCAGAAAT 1198
QY 481 GGCATGGATACCCCGTGAGTTACCCGCGGCGCGCTCGTTCAATTCAGTTTGTGAAC 540
DB 1199 GGCATGGATACCCCGTGAGTTACCCGCGGCGCGCTCGTTCAATTCAGTTTGTGAAC 1252
QY 541 CCGTGGAGGACGGGACAGCTCGCGGTGCAATGTGTTTACAGCGTATGAGCAGATGA 600
DB 1253 CCGTGGAGGACGGGACAGCTCGCGGTGCAATGTGTTTACAGCGTATGAGCAGATGA 1312

QY 601 AGATGCTCGACACGCTGCAGAACACGACGCTAGATTAACTAGAAAGATAATCATATTG 660
DB 1313 AGATGCTCGACACGCTGCAGAACACGACGCTAGATTAACTAGAAAGATAATCATATTG 1372
QY 661 TGACGCTACGTTTAAAGATAATCATGCTGTTAAATTTGACGCGATGGATCC 707
DB 1373 TGACGCTACGTTTAAAGATAATCATGCTGTTAAATTTGACGCGATGGATCC 1419
RESULT 9
ADP70043
ID ADP70043 standard; DNA; 48502 BP.
XX AC ADP70043;
XX
XX 26-AUG-2004 (first entry)
XX
XX Nucleic acid detection-related bacteriophage lambda DNA SeqID6.
XX
XX nucleic acid detection; nucleic acid amplification; microchip;
XX isothermal; chip substrate; ds.
XX
XX Bacteriophage lambda.
XX
XX JP2004154008-A.
XX
XX 03-JUN-2004.
XX
XX 01-NOV-2002; 2002JP-00320330.
XX
XX 01-NOV-2002; 2002JP-00320330.
XX
XX (EIKE) EIKEN KAGAKU KK.
XX
XX WPI; 2004-445081/42.
XX
XX Detecting nucleic acid such as DNA, cDNA and RNA, involves performing
PT nucleic acid amplification in microchip under isothermal conditions, and
PT analyzing obtained amplified product on microchip.
XX
XX Example 2; SEQ ID NO 6; 49pp; Japanese.
XX
XX This invention relates to a novel method of detecting nucleic acid which
CC comprises performing nucleic acid amplification in a microchip under
CC isothermal conditions, and analysing the obtained amplified product on
CC the microchip. The method is useful for detecting nucleic acid and
CC enables performing a nucleic acid amplification reaction and its analysis
CC continuously on the microchip without affecting the chip substrate. The
CC method enables simple, reliable and cost-effective amplification,
CC analysis and detection of nucleic acid in a microchip. The invention
CC enables detection of trace amounts of nucleic acid with amplification
CC efficiency. The present sequence is that of a bacteriophage lambda DNA
CC sequence which was used in the exemplification of the invention.
XX
XX Sequence 48502 BP; 12334 A; 11362 C; 12820 G; 11986 T; 0 U; 0 Other;
SQ
Query Match 82.8%; Score 585.4; DB 12; Length 48502;
Best Local Similarity 99.8%; Pred. No. 8e-174;
Matches 586; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGGTCTAGTCATCGCCCAAGC 104
DB 3055 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGGTCTAGTCATCGCCCAAGC 3114
QY 105 TGGCGCTATCTGGGCATCGGGAGGAGAGCCGCTGCTTTTCCGCGAGGTGAGCG 164
DB 3115 TGGCGCTATCTGGGCATCGGGAGGAGAGCCGCTGCTTTTCCGCGAGGTGAGCG 3174
QY 165 GCATGGAAGAGTGTTCGCCGAGGATGACTGCTGCTGATTGACGTTGAGGAGAAACGACG 224
DB 3175 GCATGGAAGAGTGTTCGCCGAGGATGACTGCTGCTGATTGAGGAGAAACGACG 3234

QY 225 TTTACCATGATGATTTCGGGAAGGTGTGGCCATGCAGCCCTTTAAACGGTGAACTGTTTCGTT 284
DB |||||
DB 3235 TTTACCATGATGATTTCGGGAAGGTGTGGCCATGCAGCCCTTTAAACGGTGAACTGTTTCGTT 3294
QY 285 CAGGCCACCTGGGATACCAAGTTTCGTGCGCGCTTTTCGGGACACAGTTCGGATGGTCAGC 344
DB |||||
DB 3295 CAGGCCACCTGGGATACCAAGTTTCGTGCGCGCTTTTCGGGACACAGTTCGGATGGTCAGC 3354
QY 345 CCGAAGCGCATACGACACCGAACAATACCGGCGACAGCCGGAACCTGCCGTGCGGTGTG 404
DB |||||
DB 3355 CCGAAGCGCATACGACACCGAACAATACCGGCGACAGCCGGAACCTGCCGTGCGGTGTG 3414
QY 405 CAGATTAAATCAGACCGGTGCGCGCTGGGATATTAGTTCAGCAGGACCGGGTATCTCTGGC 464
DB |||||
DB 3415 CAGATTAAATCAGACCGGTGCGCGCTGGGATATTAGTTCAGCAGGACCGGGTATCTCTGGC 3474
QY 465 TGGATGCGCGAGAAATGGACATGATACCCCGTGTAGTTTACCGCGCGCGCGCTCGTTC 524
DB |||||
DB 3475 TGGATGCGCGAGAAATGGACATGATACCCCGTGTAGTTTACCGCGCGCGCGCTCGTTC 3534
QY 525 ATTACGTTTTTGAACCCGTGGAGGACGGCGACACTCGCGGTGCAAAATGTTTTACAGC 584
DB |||||
DB 3535 ATTACGTTTTTGAACCCGTGGAGGACGGCGACACTCGCGGTGCAAAATGTTTTACAGC 3594
QY 585 GTGATGGAGCAGATGAAGATGCTTCGACACCGCTGCAGAAACACGCGAGCT 631
DB |||||
DB 3595 GTGATGGAGCAGATGAAGATGCTTCGACACCGCTGCAGAAACACGCGAGCT 3641

RESULT 10

AAS44733

ID AAS44733 standard; DNA; 4164 BP.

XX AC AAS44733;

XX DT

XX 18-DEC-2001 (first entry)

XX Bacteriophage lambda full-length polynucleotide sequence #158.

XX Mammal; human; rhesus monkey; baker's yeast; fission yeast; Norway rat;
KW mouse; Chinese hamster; African clawed frog; fruit fly; dog; leukaemia;
KW cancer; lymphoma; neuroblastoma; autoimmune disorder; cell proliferation;
KW nervous system disorder; inflammatory disorder; cell differentiation; ds;
KW angiogenesis; stem cell growth factor; activin; inhibin; cartilage; burn;
KW genetic disorder; bone regeneration; tendon; ligament; tissue repair;
KW cytoskeletal; anirrhematic; antiarthritic; vulnery; antiinflammatory;
KW antibacterial; immunosuppressive; vasotropic; antiparkinsonian;
KW neuroprotective; osteopathic; antidiabetic; antiasthmatic; antiallergic;
KW immunostimulant; analgesic; gene therapy.

XX Bacteriophage lambda.
XX WO200164834-A2.

XX PD 07-SEP-2001.

XX PF 26-FEB-2001; 2001WO-US004926.

XX PR 28-FEB-2000; 2000US-00515126.

XX PR 18-MAY-2000; 2000US-00577409.

XX PR 17-JUN-2000; 2000US-00597707.

XX PR 14-JUL-2000; 2000US-00616807.

XX PR 19-SEP-2000; 2000US-00664641.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao Q, Ren F;

XX Xue AJ, Yang Y, Wehrman T, Wang J, Ma Y, Wang D, Chen R, Xu C;

XX Drmanac R;

XX WPI; 2001-589862/66.

XX P-PSDB; AAU27833.

XX

PT Novel polypeptides and nucleic acids obtained from cDNA libraries
PT prepared from various human tissues, for diagnosis, treatment of cancer,
PT neurological, inflammatory disorders and for use in arrays for detection.

XX Claim 1; SEQ ID NO 158; 153pp; English.

XX Sequences AAS44576-AAS44919 represent full-length polynucleotides and
CC contig polynucleotides encoding polypeptides of the invention. The DNA
CC and protein sequences are useful for the treatment, diagnosis and
CC prevention of various types of disorder in a mammalian subject such as a
CC human, dog, monkey, mouse, hamster or rat. The disorders include cancers
CC such as leukaemia, lymphoma and neuroblastoma, autoimmune disorders such
CC as multiple sclerosis, connective tissue disease, rheumatoid arthritis,
CC diabetes mellitus, allergic rhinitis, asthma and eczema, nervous system
CC disorders such as Parkinson's disease, Alzheimer's disease, Huntington's
CC chorea, amyotrophic lateral sclerosis, spinal muscular atrophy and
CC Wernicke disease, inflammatory disorders such as nephritis, Crohn's
CC disease, ischaemia-reperfusion injury, shock, sepsis and inflammatory
CC bowel disease. The sequences exhibit activity relating to angiogenesis,
CC cell proliferation, cell differentiation, stem cell growth factor,
CC activin or inhibin. Therefore, they can be used to manipulate stem cells
CC in culture to give rise to neuroepithelial cells that can be used to
CC augment or replace cells damaged by illness, accidental damage or genetic
CC disorders. The sequences may also be used for regeneration of bone,
CC cartilage, tendons and ligaments and in tissue repair and burn healing.
CC Note: Some sequences for this patent did not form part of the printed
CC specification, but were obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 4164 BP; 906 A; 1123 C; 1307 G; 828 T; 0 U; 0 Other;

QY Query Match 82.1%; Score 580.6; DB 4; Length 4164;

DB Best Local Similarity 99.3%; Pred No. 7.8e-173;

DB Matches 583; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 45 ATCTAGTGCATCAGGATCATATCGTCGGGTCTTTTCCGGCTCAGTCATCGCCCAAGC 104

DB 178 ATCCAGTGCATCAGGATCATATCGTCGGGTCTTTTCCGGCTCAGTCATCGCCCAAGC 237

QY 105 TGGCGCTATCTGGGCATCGGGAGGAAGACCCGCTGCTTTCCGGGAGGTTCGAGG 164

DB 238 TGGCGCTATCTGGGCATCGGGAGGAAGACCCGCTGCTTTTCCGGGAGGTTCGAGG 297

QY 165 GCATGGAAGAGCTTTTCCGAGGATCACTGCTCTCATTTGAGTTGAGCGAAACGACG 224

DB 298 GCATGGAAGAGCTTTTCCGAGGATCACTGCTCTCATTTGAGTTGAGCGAAACGACG 357

QY 225 TTTACCATGATGATTTCGGGAAGGTGTGGCCATGACAGCTTTTAAACGGTGAATCTGTT 284

DB 358 TTTACCATGATGATTTCGGGAAGGTGTGGCCATGACAGCTTTTAAACGGTGAATCTGTT 417

QY 285 CAGGCCACCTGGGATACCAAGTTTCGTGCGGTCTTTTCCGGACACAGTTCGGATGGTCAGC 344

DB 418 CAGGCCACCTGGGATACCAAGTTTCGTGCGGTCTTTTCCGGACACAGTTCGGATGGTCAGC 477

QY 345 CCGAAGCGCATCAGCAACCCGAAACAATACCGCGGACAGCCGGAACCTGCCGTGCGGTGTG 404

DB 478 CCGAAGCGCATCAGCAACCCGAAACAATACCGCGGACAGCCGGAACCTGCCGTGCGGTGTG 537

QY 405 CAGATTAAATGACAGCGGTGCGCGCTGGGATATTAGTTCAGCAGGAGGACGGGTATCTCTGGC 464

DB 538 CAGATTAAATGACAGCGGTGCGCGCTGGGATATTAGTTCAGCAGGAGGACGGGTATCTCTGGC 597

QY 465 TGGATGCGCGAGAAATGGACATGATACCCGTGTAGTTTACCCGGCGCGCGCTCGTTC 524

DB 598 TATATGCCGCGAGAAATGGACATGGATACCCCGTGTAGTTTACCCGGCGCGCGCTCGTTC 657

QY 525 ATTACAGTTTTTGAACCCGTGGAGGACGGCGAGACTCGCGGTGCAAAATGTTTTACAGC 584

DB 658 ATTACAGTTTTTGAACCCGTGGAGGACGGCGAGACTCGCGGTGCAAAATGTTTTACAGC 717

QY 585 GTGATGGAGCAGATGAAGATGCTTCGACAGCTGCGAGAACACGCGAGCT 631

DB |||||

QY |||||

DB |||||

QY |||||

DB |||||

QY |||||

DB |||||

QY |||||

DB |||||

QY |||||

DB |||||

QY |||||

DB |||||

QY |||||

DB |||||

QY |||||

DB |||||

QY |||||

DB |||||

QY |||||

DB |||||

QY |||||

DB |||||

QY |||||

DB |||||


```
Db      718 GTGATGGAGCAGATAAAGATGCTCGACACGCTGCAGAAACGCGAGCT 764
RESULT 11
ABS78924/c
ID      ABS78924 standard; DNA; 26565 BP.
XX
AC      ABS78924;
XX
DT      17-DEC-2002 (first entry)
XX
DE      E. coli CFT073 genomic sequence #91.
XX
XX      Pathogenic; Escherichia coli CFT073 infection; livestock; pyelonephritis;
KW      urinary tract infection; open reading frame; ORF; uropathogenic;
KW      antibacterial; aropathic; nephrotropic; gene; ds.
XX
XX      Escherichia coli.
OS
XX
PN      WO200259320-A2.
XX
PD      01-AUG-2002.
XX
PF      19-OCT-2001; 2001WO-US046833.
XX
PR      19-OCT-2000; 2000US-0242412P.
XX
XX      (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
XX      Blattner FR, Welch RA, Burland VD;
XX
XX      WPI; 2002-691532/74.
XX
XX      New DNA sequences of the pathogenic Escherichia coli CFT073 strain,
PT      useful for preventing or treating E. coli CFT073 infection in humans or
PT      livestock.
XX
XX      Claim 1; Page 276-291; 765pp; English.
XX
XX      The present invention relates to polynucleotide sequences from the genome
CC      of the pathogenic Escherichia coli strain CFT073. Almost all the
CC      sequences present in E. coli CFT073 are absent in the previously
CC      sequenced laboratory strain K-12. The polynucleotide sequences of the
CC      invention are useful for preventing, diagnosing or treating E. coli
CC      CFT073 infection in humans or livestock. The polynucleotide sequences are
CC      useful for preventing urinary tract infections and pyelonephritis.
CC      Likewise, the polypeptides encoded by the different open reading frames
CC      (ORF1-5) are useful for generating a vaccine against uropathogenic E.
CC      coli strains. ABS78834-ABS79085 represent genomic sequences from E. coli
CC      strain CFT073
XX
XX      Sequence 26565 BP; 7042 A; 6727 C; 5778 G; 7016 T; 0 U; 2 Other;
SQ
Query Match      80.1%; Score 566.2; DB 6; Length 26565;
Best Local Similarity 97.8%; Pred. No. 7.2e-168;
Matches 574; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
QY      45 ATCTAGCTGCATCAGGATCATATGTCGGGTCTTTTTCGGGCTCAGTCATCCGCCAAGC 104
Db      3221 ATCCAGCTGCATCAGGATCATATGTCGGGTCTTTTTCGGGCTCAGTCATCCGCCAAGC 3162
QY      105 TGGCGCTATCTGGGCATCGGGGAGGAAGAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 164
Db      3161 TGGCGCTATCTGGGCATCGGGGAGGAAGAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 3102
QY      165 GCATGGAAGAGTTTCCCGAGGATGACTGCTGCTGCAATTGACGTTGAGCGAAACGCGACG 224
Db      3101 GCATGGAAGAGTTTCCCGAGGAGCACTGTTGCTGCAATTGACGTTGAGCGAAACGCGACG 3042
QY      225 TTATCCATGATGATTCGGGAGGTGTGCCCATGCAGCCCTTTTAACCGTGAACTGTTTCGTT 284
Db      3041 TTACCATGATGATTCGGGAGGAGGTGTGCCCATGCATGCCTTTTAACCGTGAACTGTTTCGTT 2982
```

QY 45 ATCTAGCTGCATCAGGATCATATCGTGGGTCTTTTTCGGCTCAGTCATCGCCCAAGC 104
DB |||||
3221 ATCCAGCTGCATCAGGATCATATCGTGGGTCTTTTTCGGCTCAGTCATCGCCCAAGC 3162
QY 105 TGGCGCTATCTGGCATCGGGAGGAGAACGCCGTGCTTTTCCCGGAGGTTGAAGCG 164
DB |||||
3161 TGGCGCTATCTGGCATCGGGAGGAGAACGCCGTGCTTTTCCCGGAGGTTGAAGCG 3102
QY 165 GCATGGAGAGTTCCCGGAGGATGACTGCTGTGCAATTGACGTTGAGCGAACCACG 224
DB |||||
3101 GCATGGAGAGTTCCCGGAGGAGGACTGTGCTGATTGACGTTGAGCGAACCACG 3042
QY 225 TTTACCATGATGATTCGGGAAGTGTGGCCATCGACGCCCTTTAAACGGTGAACCTGTT 284
DB |||||
3041 TTACCATGATGATTCGGGAAGTGTGGCCATCGACGCCCTTTAAACGGTGAACCTGTT 2982
QY 285 CAGGCCACCTGGGATACCAAGTTCTGTCGGGCTTTTCCGACACAGTTCGGATGFTCAGC 344
DB |||||
2981 CAGGCCACCTGGGATACCAAGCTCGGACTGTTCCGGACACAGTTCGGATGFTCAGT 2922
QY 345 CCGAAGGCATCAGCAACCCGAACAATACCGGCGACAGCGGNACTGCCGTGCCGTGTG 404
DB |||||
2921 CCGAAGGCATCAGCAACCCGAACAATACCGGCGACAGCGGNACTGCCGTGCCGTGTG 2862
QY 405 CAGATTAAATGACAGCGGTGGCGCTGGGATATTAGCTCAGCGAGACGGGTATCTGGC 464
DB |||||
2861 CAGATTAAATGACAGCGGTGGCGCTGGGATATTAGCTCAGCGAGACGGGTATCTGGC 2802
QY 465 TGGATGCCGAGAAATGGACATGGATACCCCGTGAAGTTACCGCGGGCGCGCTCGTTTC 524
DB |||||
2801 TGGATGCCGAGAAATGGACATGGATACCCCGTGAAGTTACCCGCGGGCGCGCTCGTTTC 2742
QY 525 ATTACGTTTTTGAACCCGTGGAGGACGGCGACAGCTCGCGGTGCAAAATGTGTTACAGC 584
DB |||||
2741 ATTACGTTTTTGAACCCGTGGAGGACGGCGACAGCTCGCGGTGCAAAATGTGTTACAGC 2682
QY 585 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGCGAGCT 631
DB |||||
2681 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGCGAGCT 2635

RESULT 13
ACD19059
ID ACD19059 standard; DNA; 46819 BP.
XX
AC ACD19059;
XX
DT 27-OCT-2003 (revised)
DT 21-AUG-2003 (first entry)
XX
DE E. coli 0157 unique DNA sequence OZID_72.
XX
XX OZID; ds; acute haemorrhagic colitis; haemolytic uraemic syndrome;
KW food poisoning.
XX
XX Escherichia coli; strain 0157:H7.
XX
FN US2003023075-A1.
XX
XX 30-JAN-2003.
PD
XX 01-APR-2002; 2002US-00114170.
PF
XX 04-DEC-1998; 98US-0110955P.
PR 03-DEC-1999; 99US-00453702.
XX
XX (BLAT/) BLATTNER F R.
PA (BURL/) BURLAND V D.
PA (PERN/) PERNA N T.
PA (PLUNK/) PLUNKETT G.
PA (WELC/) WELCH R.
XX
XX Blattner FR, Burland VD, Perna NT, Plunkett G, Welch R;

XX WPI; 2003-479497/45.
DR
XX
PT New DNA sequences from Escherichia coli strain 0157:H7, useful for
PT detecting E. coli 0157:H7 in a sample, or in designing diagnostic probes
PT which can be used to distinguish strain 0157:H7 from strain K12 using
PT molecular techniques.
XX
PS Claim 16; SEQ ID NO 72; 33pp; English.
XX
CC The invention relates to an isolated DNA molecule comprising an E. coli
CC strain 0157:H7 sequence selected from a clostridial cytotoxin-like gene,
CC a urease gene cluster, a RTX toxin-like gene cluster, a locus of
CC enterocyte effacement and 2 genes from its associated lymphocytic phage
CC 933W (a putative serine/threonine kinase and a tail fibre gene). E. coli
CC 0157:H7 can cause food poisoning, specifically acute haemorrhagic colitis
CC (which can develop into haemolytic uraemic syndrome). Also included are
CC an isolated DNA molecule comprising a nucleotide sequence identical to at
CC least 25 contiguous nucleotides contained in DNA sequences selected from
CC ACD18988-ACD19242 (being 255 E.coli 0157 DNA sequences which are not
CC found in E.coli K12), a recombinant DNA construction comprising the DNA
CC above and a method for detecting E. coli 0157:H7 (ATCC 43895) in a sample
CC (or distinguishing between 0157 and K12) using a probe derived from one
CC of the 255 sequences. The DNA sequences are useful in detecting E. coli
CC 0157:H7 in a sample, for the early diagnosis of humans and livestock
CC infected with 0157:H7, and in designing diagnostic probes which can be
CC used to distinguish strain 0157:H7 from strain K12 using molecular
CC techniques. The present sequence is one of the 255 E. coli 0157:H7 DNA
CC sequence (termed OZID.1-OZID255). Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from the USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030023075 (Updated on 27-OCT-2003
CC to standardise OS field)
XX
SQ Sequence 46819 BP; 11858 A; 10855 C; 12774 G; 11317 T; 0 U; 15 Other;
Query Match 80.1%; Score 566.2; DB 9; Length 46819;
Best Local Similarity 97.8%; Pred. No. 9.6e-168;
Matches 574; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
QY 45 ATCTAGCTGCATCAGGATCATATCGTGGGTCTTTTTCGGCTCAGTCATCGCCCAAGC 104
DB |||||
14795 ATCCAGCTGCATCAGGATCATATCGTGGGTCTTTTTCGGCTCAGTCATCGCCCAAGC 14854
QY 105 TGGCGCTATCTGGCATCGGGAGGAGAACGCCGTGCTTTTCCCGGAGGTTGAAGCG 164
DB |||||
14855 TGGCGCTATCTGGCATCGGGAGGAGAACGCCGTGCTTTTCCCGGAGGTTGAAGCG 14914
QY 165 GCATGGAAAGAGTTTCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAGCGCAG 224
DB |||||
14915 GCATGGAAAGAGTTTCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAGCGCAG 14974
QY 225 TTTACCATGATGATTCGGGAAGTGTGGCCATCGACGCCCTTTAAACGGTGAACCTGTT 284
DB |||||
14975 TTTACCATGATGATTCGGGAAGTGTGGCCATCGACGCCCTTTAAACGGTGAACCTGTT 15034
QY 285 CAGGCCACCTGGGATACCAAGTTCTGTCGGGCTTTTCCGACACAGTTCGGATGFTCAGC 344
DB |||||
15035 CAGGCCACCTGGGATACCCGTCCTCGGACTGTTCCGGACACAGTTCGGATGFTCAGC 15094
QY 345 CCGAAGGCATCAGCAACCCGAACAATACCGGCGACAGCGGNACTGCCGTGCCGTGTG 404
DB |||||
15095 CCGAAGGCATCAGCAACCCGAACAATACCGGCGACAGCGGNACTGCCGTGCCGTGTG 15154
QY 405 CAGATTAAATGACAGCGGTGCCGCTGGGATATTAGCTCAGCGAGACGGGTATCTGGC 464
DB |||||
15155 CAGATTAAATGACAGCGGTGCCGCTGGGATATTAGCTCAGCGAGACGGGTATCTGGC 15214
QY 465 TGGATGCCGAGAAATGGACATGGATACCCCGTGAAGTTACCCGGGGCGCGCTCGTTTC 524
DB |||||
15215 TGGATGCCGAGAAATGGACATGGATACCCCGGAGTTACCCGGGGCGCGCTCGTTTC 15274
QY 525 ATTACGTTTTTGAACCCCGTGGAGGACGGGCGAGACTCGCGGTGCAAAATGTGTTTACAGC 584

CC found in E.coli K12), a recombinant DNA construction comprising the DNA
CC above and a method for detecting E. coli 0157:H7 (ATCC 43895) in a sample
CC (or distinguishing between 0157 and K12) using a probe derived from one
CC of the 255 sequences. The DNA sequences are useful in detecting E. coli
CC 0157:H7 in a sample, for the early diagnosis of humans and livestock
CC infected with 0157:H7, and in designing diagnostic probes which can be
CC used to distinguish strain 0157:H7 from strain K12 using molecular
CC techniques. The present sequence is one of the 255 E. coli 0157:H7 DNA
CC sequence (termed O2ID.1-O2ID255). Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from the USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030023075 (Updated on 27-OCT-2003
CC to standardise OS field)
XX
SQ Sequence 22306 BP; 4744 A; 6842 C; 5512 G; 5192 T; 0 U; 16 Other;

Query Match 79.4%; Score 561.4; DB 9; Length 22306;
Best Local Similarity 97.3%; Pred. No. 2.2e-166;
Matches 571; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 45 ATCTAGTCGATCAGGATCATATCGTCGGTCTTTTTCGGCTCAGTCATCGCCCAAGC 104
Db |||||
QY 17193 ATCAGTTGATCAGGATCATATCGTCGGTCTTTTTCGGCTCAGTCATCGCCCAAGC 17134
Db |||||
QY 105 TGGCGCTATCTGGGCGATCGGGGAGGAAGCCCGTGCCTTTTCCCGGAGGTTGAAGCG 164
Db |||||
QY 17133 TGGCGCTATCTGGGCGATCGGGGAGGAAGCCCGTGCCTTTTCCCGGAGGTTGAAGCG 17074
Db |||||
QY 165 GCATGGAAGAGTTTCCGAGGATGACTGCTGTCATTGACGTTGAGCGAAGAACGCACG 224
Db |||||
QY 17073 GCATGGAAGAGTTTCCGAGGATGACTGCTGTCATTGACGTTGAGCGAAGAACGCACG 17014
Db |||||
QY 225 TTTTACCATGATGATTCGGGAAGGTGGCCATCGACGCTTTTAAACGGTGAACCTGTTTCGTT 284
Db |||||
QY 17013 TTTTACCATGATGATTCGGGAAGGTGGCCATCGACGCTTTTAAACGGTGAACCTGTTTCGTT 16954
Db |||||
QY 285 CAGGCCACCTGGGATACCAAGTTTCGTCGGCTTTTCCGGACACAGTTCGGATGGTTCAGC 344
Db |||||
QY 16953 CAGGCCACCTGGGATACCAAGTTTCGTCGGCTTTTCCGGACACAGTTCGGATGGTTCAGC 16894
Db |||||
QY 345 CCGAAGCGCATCAGCAACCCGGAACAATACCGGCGACAGCCGGAACCTGCCGTGCCGTGTG 404
Db |||||
QY 16893 CCGAAGCGCATCAGCAACCCGGAACAATACCGGCGACAGCCGGAACCTGCCGTGCCGTGTG 16834
Db |||||
QY 405 CAGATTAAATGACAGCGGTGGCGCTGGGATATTACGTAGCGAGGACGGGTATCCTGGC 464
Db |||||
QY 16833 CAGATTAAATGACAGCGGTGGCGCTGGGATATTACGTAGCGAGGACGGGTATCCTGGC 16774
Db |||||
QY 465 TGGATGCCGAGAAATGGACATGGATACCCCGTGGTATACCCGGCGGGCGCGCTCGTTC 524
Db |||||
QY 16773 TGGATGCCGAGAAATGGACATGGATACCCCGTGGTATACCCGGCGGGCGCGCTCGTTC 16714
Db |||||
QY 525 ATTACGTTTTTGAACCCGTGGAGGACGGGCGACACTCGCGGTGCAATGTGTTTACAGC 584
Db |||||
QY 16713 ATTACGTTTTTGAACCCGTGGAGGACGGGCGACACTCGCGGTGCAATGTGTTTACAGC 16654
Db |||||
QY 585 GTGATGGAGCAGATGAAGATGCTCGACACGCTCGAGAACACGCGAGCT 631
Db |||||
QY 16653 GTGATGGAGCAGATGAAGATGCTCGACACGCTCGAGAACACGCGAGCT 16607
Db |||||

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 5, 2005, 13:37:08 ; Search time 2963.77 Seconds
(without alignments)

11558.877 Million cell updates/sec

Title: US-10-826-523-40

Perfect score: 707

Sequence: 1 ggaatccatcgctcaattt.....aaaattgaacgatggatcc 707

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hg.*

3: gb_in.*

4: gb_on.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	585.4	82.8	20067	12 XXU02425	U02425 Cloning vec
2	585.4	82.8	20125	12 XXU02426	U02426 Cloning vec
3	585.4	82.8	42529	12 CVU39284	U39284 Cloning vec
4	585.4	82.8	42530	12 CVU39285	U39285 Cloning vec
5	585.4	82.8	42531	12 CVU39286	U39286 Cloning vec
6	585.4	82.8	42704	12 CVU37692	U37692 Cloning vec
7	585.4	82.8	48502	7 LAMCG	J02459 Bacterioph
8	574.6	81.3	5969	12 AF434923	AF434923 Expressio
9	574.6	81.3	6182	12 AF434925	AF434925 Expressio
10	574.6	81.3	6182	12 AF434926	AF434926 Expressio
11	574.6	81.3	6182	12 AF434927	AF434927 Expressio
12	574.6	81.3	6182	12 AF434931	AF434931 Expressio
13	574.6	81.3	6182	12 AF434932	AF434932 Expressio
14	574.6	81.3	7421	12 AF434924	AF434924 Expressio
15	574.6	81.3	7634	12 AF434928	AF434928 Expressio
16	574.6	81.3	7634	12 AF434929	AF434929 Expressio
17	574.6	81.3	7634	12 AF434930	AF434930 Expressio
18	574.6	81.3	7634	12 AF434933	AF434933 Expressio
19	574.6	81.3	7634	12 AF434934	AF434934 Expressio

20	566.2	80.1	14164	1	AE005330	Escherich
c 21	566.2	80.1	26565	6	AX702505	Sequence
22	566.2	80.1	46819	6	AR204176	Sequence
c 23	566.2	80.1	305325	1	AE016765	Escherich
24	564.6	79.9	46897	6	BD184772	Nucleic a
c 25	564.6	79.9	222605	1	AP002555	Escherich
26	564.6	79.9	222605	6	AR204355	Sequence
c 27	561.4	79.4	22306	6	AR204355	Sequence
c 28	561.4	79.4	91772	6	BD184765	Nucleic a
c 29	561.4	79.4	286485	1	AP002557	Escherich
c 30	395.8	56.0	46375	7	AF064539	Bacterioph
c 31	395.8	56.0	46375	2	AC101160	Mus muscu
c 32	395.8	56.0	179871	2	AC024602	Homo sapi
c 33	395.8	56.0	194028	2	AC111077	Homo sapi
c 34	395.8	56.0	206486	2	AC074218	Mus muscu
c 35	362.2	51.2	228516	2	AC008594	Homo sapi
c 36	343.2	48.5	68028	2	AC101473	Mus muscu
c 37	329.8	46.6	66450	2	AC060821	Homo sapi
c 38	320.4	45.3	63579	2	AC100018	Mus muscu
c 39	320.2	45.3	63690	2	AC125419	Homo sapi
c 40	231.8	32.8	29466	1	AE008818	Salmonell
c 41	223	31.5	8134	7	PH2HEADTL	M81255 Bacterioph
c 42	223	31.5	13264	1	AE005368	Escherich
c 43	223	31.5	13386	1	AE015100	Shigella
c 44	223	31.5	34063	6	AR204200	Sequence
c 45	223	31.5	44029	6	BD184774	Nucleic a

ALIGNMENTS

RESULT 1
XXU02425
LOCUS
DEFINITION Cloning vector lambda EMBL3, left arm.
ACCESSION U02425
VERSION U02425.1 GI:413791
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 20067)
AUTHORS Frischauf,A.M., Lehrach,H., Pouetka,A. and Murray,N.
TITLE Lambda replacement vectors carrying polylinker sequences
JOURNAL J. Mol. Biol. 170 (4), 827-842 (1983)
MEDLINE 84064856
PUBMED 6315951
REFERENCE 2 (bases 1 to 20067)
AUTHORS Kitts,P.A.
TITLE CLONTECH Vectors On Disc version 1.3
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 20067)
AUTHORS Kitts,P.A.
TITLE Direct Submission
JOURNAL Submitted (07-Oct-1993) Paul A. Kitts, CLONTECH Laboratories, Inc., 1020 East Meadow Circle, Palo Alto, CA 94303, USA
COMMENT This sequence has been compiled from information in the sequence databases, published literature and other sources. If you suspect there is an error in this sequence, please contact CLONTECH's Technical Service Department at (415) 424-8222 or (800) 662-2566, extension 3 or E-mail TECH@CLONTECH.COM.

FEATURES
source
1. .20067
/organism="Cloning vector lambda EMBL3"
/mol_type="genomic DNA"
/db_xref="taxon:31783"

ORIGIN

Query Match 82.8%; Score 585.4; DB 12; Length 20067;
Best Local Similarity 99.8%; Pred. No. 2.3e-148;
Matches 586; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 45 ATCTAGTCGATCAGGATCATATCGTGGGTCTTTTTCGGGCTCAGTCATCCGCCAAGC 104

Query Match 82.8%; Score 585.4; DB 12; Length 42529;
Best Local Similarity 99.8%; Pred. No. 2.5e-148;

gene	complement(19507..20760) /gene="lacY"	Matches 586; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
CDS	complement(19507..20760) /gene="lacY" /codon_start=1 /transl_table=11 /product="lactose permease" /protein_id="AAC53651.1" /db_xref="GI:1066309" /translation="MYLKNTNFMFGLPFFFFFFIMGAYFPFPFPIWLHDIHISKSD TGIIFAAISLFLFQGLSDGLKRLKYLWIIITGLMFMFAFPFIFIFGLLQYN ILVGSIVGGYVGLFCFNAGAPAVEATEKYSRRSNFEFGARMFGVGCWALCASIVGI MFTINQVFWLGGSCALILAVLFFAKTDAPSATVANAVAGHSAFSLKLALEIFR QPKLFLSLVIVGVSCTVDVDOOFANFETSFPATGEQGTFRVGYVITWGLNLASIM FPAPLINRIGGKNALLAGTINSVRILIGSSPATSALEVILKTLHMFEVFPFLVGCFF KYITQFEVFAFSATILVCFCKQALMIFMSVLGNMYESIGFQAYILVLGLVALGF TLISVFTLGGPFLSLLRQNEVA" complement(20812..23859) /gene="lacZ" /product="beta-galactosidase" /protein_id="AAC53652.1" /db_xref="GI:1066310" /translation="VLQRRDWNPGVTQLNRLAAHPPFASWRNSBEATDRPSQQLR SLNGWRFAFPFAPEAFSECDLPEADTVVPSNQWQHGYPADIYTNVTPITVN PPFVPTNPFGCSLTFNVDSESLQSGTIRIIFDVNSAFHLMCNGRWVGYGQDSRLP SEFDLGAFLRAGNRLAVMLRWSGYLEDDQWRMSGIFRDVSLHLKPTTQISDFH VATRENDLSEPAVLEAVOMCGELRYLRTVLSMOGETQVASTAPRGEIILDERGG YADRVLINVENPKLWSAEIPNLRYAVVELHTADGTLIEAEACDVGFREVLNELL LLMGKLLINGVRHHEHPLHUGQDEQTNVDIILMKQNNFNAVRCSHYPNPLMYT LCDRYGLYVDEANIEITHGVPNRLTDDPRWLPAMSERVTMVQDRHPSVIIWSL GNESGHAHDALYRIKSVDPSPVQYEGGGADTTATDIICPMYARVDEDDPFPAPV KWSIKKSLSPGTRPLILCEYAHAMGNSLGFAYKWOAFPRVYRQLOGGVFWDVQDS LIXYDNGNPMWAVGDDPTNDROFCWNLGYEADRTPHALTEAKHQOQPEFQSL AQGLVTRVVQPNATSEAGHSIAWQWRLAENLSVTLPAASHAIPHLTSEMDFC IELGNRWQFQSGFLSQWIGDKKQLLTPLRDQFTRAPLDNDIIVSEATRIDNPAW VERWKAALHYQAEALLOCTADTLADAVLITTAHAWHQOKTLFI SRKTYRIDGSGQM AITVDVEASDTPHARIGLNCOLAQVAERVNWLGLPOENYPRDLTAACFRDWDLP SDMYTPVFPSENGLCRTRELNGYBHQWGDFOENISRYSOQLMETSHRLLHAE GTWLNDGFMHGGDDSWSPSBAEQLSAGRYHQLVWCQK" complement(23875..23958) /note="multiple cloning site" /notes="tetR terminator" /notes="rrnBT1 terminator" 24476..25666 /gene="tetR" 24476..25666 /gene="tetR" /codon_start=1 /transl_table=11 /product="tetracycline resistance protein" /protein_id="AAC53653.1" /db_xref="GI:1066311" /translation="MKSNNALIVLGTVTLDVAVGILGVMPVLGGLLRDIIVHSDSIASH YGVLLALYALMQFLCAPVLGALSDFRREVLLASLLGATIDYAIMATTPVLMWLYAG RIVAGITGATGAVAGYIADITDGEDRARHFGLMSACFGVGMVAGPVAGGLGALSILH APFLAAVLNGLNLLGQFLMQESHKGERPMLRFPNVPVSPFRWARGMTIVAALMTV FFIMQLGVQVPALWTFEGEDRFRWNSATMIGLSLAVFGILHALAQAPVVGPAKFRGE KQAILAGAAALGYVLLAFATRGWAFPIILLIAGSGGIGMPALQALQMSQVDDDDHQG QLQSGAAATLSLTSITGLPITVTAIYAASASTWNLGLAWI VGAALYLVCLPALRRGAWSR ATST"	45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGGCTAGTCATCGCCCAAGC 104 3055 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGGCTAGTCATCGCCCAAGC 3114 105 TGGCGCTATCTGGGCATCGGGGAGGAAGAGCCGTCCTTTTTCGGCGGAGGTTGAAGCG 164 3115 TGGCGCTATCTGGGCATCGGGGAGGAAGAGCCGTCCTTTTTCGGCGGAGGTTGAAGCG 3174 165 GCATGAAAAGAGTTTCCGAGGATGACTGCTGCTGCATTCAGCTTCCGAGGAAAACGACG 284 3175 GCATGAAAAGAGTTTCCGAGGATGACTGCTGCTGCATTCAGCTTCCGAGGAAAACGACG 3234 225 TTATACCATGATGATTCGGGAAGGTGGCCATGACGACGCTTTTAAACGGTGAACCTGTTTTCGTT 284 3235 TTATACCATGATGATTCGGGAAGGTGGCCATGACGACGCTTTTAAACGGTGAACCTGTTTTCGTT 3294 285 CAGGCCACCTGGGATACCACTGTCGTCGGGCTTTTTCGGGACACAGTTCGGGATGGTCAGC 344 3295 CAGGCCACCTGGGATACCACTGTCGTCGGGCTTTTTCGGGACACAGTTCGGGATGGTCAGC 3354 345 CCGAAGCGCATGACGAAACCGGAAACAATACCGGGGACAGCGGAACTGCCGTGCCGTGTG 404 3355 CCGAAGCGCATGACGAAACCGGAAACAATACCGGGGACAGCGGAACTGCCGTGCCGTGTG 3414 405 CAGATTAAATGACAGCGGTGCGGCTGGGATATTACGTACGAGGAGGACGGGTATCTTGGC 464 3415 CAGATTAAATGACAGCGGTGCGGCTGGGATATTACGTACGAGGAGGACGGGTATCTTGGC 3474 465 TGGATCCCGCAGAAATGAGCATGATACCCCGTGGTGTACCCCGCGGGCGCGCTCGTTC 524 3475 TGGATCCCGCAGAAATGAGCATGATACCCCGTGGTGTACCCCGCGGGCGCGCTCGTTC 3534 525 ATTACAGTTTTTGAACCCCTGGAGGACGGGACAGTCTCGCGGTGCAAAATGTGTTTACAGC 584 3535 ATTACAGTTTTTGAACCCCTGGAGGACGGGACAGTCTCGCGGTGCAAAATGTGTTTACAGC 3594 585 GTGATCGAGCAGATGAAGATGCTCGACACGCTGACAGAACACGCGCT 631 3595 GTGATCGAGCAGATGAAGATGCTCGACACGCTGACAGAACACGCGCT 3641
misc_feature	RESULT 5 LOCUS CUU39286 42531 bp DNA linear SYN 13-APR-1996 Cloning vector TLF97-3, phage lambda lacZ translational fusion vector, complete sequence.	
terminator	ACCESSION U39286 VERSION U39286.1 GI:1066312 KEYWORDS Cloning vector TLF97-3 SOURCE Cloning vector TLF97-3 ORGANISM Cloning vector TLF97-3 Other sequences; artificial sequences; vectors.	
gene	REFERENCE 1 (bases 1 to 42531) AUTHORS St Pierre,R. and Linn,T. TITLE A refined-copy system for the in vitro construction of single-copy transcriptional or translational fusions to lacZ JOURNAL Gene 169 (1), 65-68 (1996) MEDLINE 96186904 PUBMED 863751	
CDS	REFERENCE 2 (bases 1 to 42531) AUTHORS St Pierre,R. TITLE Direct Submission JOURNAL Submitted (24-OCT-1995) Thomas Linn, Microbiology and Immunology, University of Western Ontario, London, Ontario N6A5C1, Canada FEATURES Location/Qualifiers 1..42531 /organism="Cloning vector TLF97-3" /mol_type="genomic DNA" /db_xref="taxon:43842" complement(19404..19489) /note="rrnBT2 terminator"	
misc_feature	33603..35451 /note="imm21 substitution"	
ORIGIN	Query Match 82.8%; Score 585.4; DB 12; Length 42530; Best Local Similarity 99.8%; Pred. No. 2.5e-148;	


```
removed by an uncharacterized mutation"
5505
/misc_feature
/notes="BamHI site normally present at this site has been
removed by an uncharacterized mutation"
complement(19404..19499)
/notes="rrnB72 terminator"
gene
/genes="lacY"
complement(19507..20760)
CDS
/genes="lacY"
/codon_start=1
/transl_table=11
/product="lactose permease"
/protein_id="AAC53645.1"
/db_xref="GI:1051182"
/translation="MYLKNTEWMEGLPEFFYFFITMGAYEPFPIWLHDIHISKSD
TGLIFASISLFLPQLGSLDKLGRKYLMLIITGMLVMPAFFIFIGPILLYN
ILVGSVGGVIGFCFNAGAPAEAFIEKVSRSNFEFGARMGFCGVALCAGISVI
MFTINQFVWLGSCALILAVLLFAKTDAPSATVANAGANHSFAFSLKALELFR
QPKLWFLSVLVGSCVYDVFQDFANFTSPFATGEQGTGVFGYVTTMGLINASIM
FPAPLINRIGGNALLAGTINSVRIIGSSPATSALEVILKTHMFVPPFLVGCFF
KYITSQEVRFESATVILVCFEKLAMIFMSVLAGNMYESIGFQGYLVILVALGF
TLISVTLGFGFLSLRLRQNEVA"
complement(20812..23886)
/genes="lacZ"
complement(20812..23886)
/genes="lacZ"
/notes="EcoRI site normally present at the end of lacZ has
been removed by an uncharacterized mutation"
/codon_start=1
/transl_table=11
/product="beta-galactosidase"
/protein_id="AAC53646.1"
/db_xref="GI:1051183"
/translation="MTMITDSLAVLQRRDWNPGVTQLRLAAHPFPASWRNSEAR
TDRPSQLSLNGEMFAPFAPEAPESWLECDLPEADTVVVSNNQMHGYDAPIYT
NVYPTVNPPTPGCYSLTFNVDESMLQEGQTRIIFDVNSAFHLCWNGRW
GYQDRLPEEFDLAFRAGENLAVMLRWSGSLYLEDQDMWRMSGI PRDVSLLHK
PTQISDFHVAFTENDFSRAVLEAEVQMGELRDYLRVTSLWQGETOVASGTAPFG
GEILDERGGVADYRLNVENPKLSAEI PNLYRAVVELHTADGTLEAEACDVGR
EVRLEGLLNLKPLIRGVNRHEHPHGLHQVDEQTMVQDILLMKQNFNARCRSH
YNPHLYLLCDRIGLYVDEANIETHGMVPMNRLTDDPRWLPAMSERVTRMVQRDN
HPSVIVISLNGESGHGANIDALRWIKSVDPSPVQYEGGADTTA2D1ICPMYARVD
EDQFPAPKWSIKKWSLSPGETRPLILCEYAHAMNLSGSGFAKYQAQFQYPRLOCG
FVWDVQDSILKYDENGPMWSAYGGDPTNDPQFCMGLVPADTRPAPLATEAKHQ
QOFQPLSGQTLTYVSEYLFPSHDNELLHMVALDKPLAGEVPLDVAPOGKOLIE
LPFLPESAGQWLTVRVQPNATASEAGHSANQWRLEAENLSVTLPASHALIPH
LITSEMDFCIELGNKRWQFNQSGFLSQMWIGDKQLLTPLRDQTRAPLDNDIGVSE
ATRIDPNAWYERKAAHQYQAEALIQCTADTLADAVLITTAHAWQHQKTLFISRT
YRIDGQMAITVDVEVADTPHAPRIGLNCQLAQVAERVNMWLGFPQBNYPRDLTAA
CFDRWDLPLSDMYTPYVFPSENGLRCTRELNGPHQWRGDFQFNLSRYSQOQLMETS
HRLLHABEGTWNIDGFMHGIIGDDSDWSPSVSAEFLSNAGRHYHQLVNCQK"
complement(24049..24132)
/notes="multiple cloning site"
complement(24148..24316)
/notes="rrnB71 terminator"
gene
24650..25840
/genes="tetR"
24650..25840
/genes="tetR"
/notes="Description: tetracycline resistance gene;
tetracycline resistance protein"
/codon_start=1
/transl_table=11
/protein_id="AAC53647.1"
/db_xref="GI:1051184"
/translation="MKSNNALIVILGTVTLDAVIGLVMPVLPLGLLRDIIVHSDSIASH
YGVLLALYALMQFLCARFRFRFVLILASLIGATIDYIMATTPVLMWLYAG
RIVAGTAVGAGVAGVADIOTGEDRARRHGLMSACFGVGVMPVAGGLLCAILH
APFLAAVNLGNLLGCTLMQESHKERRRPMELAFNVPVSSFRWARGMTIVAALMTV
FFTMQLVGQVPALWTFGEDRFRNSATMIGLSLAVFGLIHALAQAFVTPGPKRGE
KQAITAGAADALGVLLPATRWAFPIIMILLASGGIGMPALQGLSRVDRDDHQG
QCGSLAALTSLSITGLPIVTAIYAASASTWNLAWIYGAALYLVCLFLPRLRRGASR
```

```
ATST"
31747
/misc_feature
/notes="EcoRI site normally present at this site has been
removed by an uncharacterized mutation"
33777..35625
/notes="imm21 substitution"
39168
/misc_feature
/notes="EcoRI site normally present at this site has been
removed by an uncharacterized mutation"
ORIGIN
Query Match 82.8%; Score 585.4; DB 12; Length 42704;
Best Local Similarity 99.8%; Pred. No. 2.5e-148;
Matches 586; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 45 ATCTAGCTCATCAGGATCATATCGTCGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 104
DB 3055 ATCCAGCTGCATCAGGATCATATCGTCGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 3114
QY 105 TGGCGCTATCTGGGCATCGGGGAGGAAGAGCCGTCCTTTTCCCGGAGGTTGAAGCG 164
DB 3115 TGGCGCTATCTGGGCATCGGGGAGGAAGAGCCGTCCTTTTCCCGGAGGTTGAAGCG 3174
QY 165 GCATGAAAAGAGTTTGGCCGAGGATGACTCTCTGCTGCAATTGACGTTGAGCGAAACCCACG 224
DB 3175 GCATGAAAAGAGTTTGGCCGAGGATGACTCTCTGCTGCAATTGACGTTGAGCGAAACCCACG 3234
QY 225 TTTACCATCATGATTCGGGAGAGTGTCGTCATGCAATGTCGTTGAGCGAAACCCACG 284
DB 3235 TTTACCATCATGATTCGGGAGAGTGTCGTCATGCAATGTCGTTGAGCGAAACCCACG 3294
QY 285 CAGGCCACCTGGGATACCAAGTTTCGTCGGCGCTTTTTCGGACACAGTTCGGGATGTCGACG 344
DB 3295 CAGGCCACCTGGGATACCAAGTTTCGTCGGCGCTTTTTCGGACACAGTTCGGGATGTCGACG 3354
QY 345 CCGAAGCGCATCAGCAACCCGAAACAATAACCGGCGACAGCGGAACTGCCGTGCGGTGTG 404
DB 3355 CCGAAGCGCATCAGCAACCCGAAACAATAACCGGCGACAGCGGAACTGCCGTGCGGTGTG 3414
QY 405 CAGATTAAATGACAGCGGTGCGCGCTGGGATATTAGTCAGCGAGGACGGGTATCTCTGCG 464
DB 3415 CAGATTAAATGACAGCGGTGCGCGCTGGGATATTAGTCAGCGAGGACGGGTATCTCTGCG 3474
QY 465 TGGATGCCCGCAGAAATGGACATGATACCCCGTAGTATTACCCGGCGGGCGCGCTCTGTC 524
DB 3475 TGGATGCCCGCAGAAATGGACATGATGATACCCCGTAGTATTACCCGGCGGGCGCGCTCTGTC 3534
QY 525 ATTACCGTTTTTGAACCCGTGAGGACCGGGCAGACTCGCGGTGCAAAATGTGTTTACAGC 584
DB 3535 ATTACCGTTTTTGAACCCGTGAGGACCGGGCAGACTCGCGGTGCAAAATGTGTTTACAGC 3594
QY 585 GTGATGGACAGATGAAGATGCTCGACACGCTGCAGAACACCGCAGCT 631
DB 3595 GTGATGGACAGATGAAGATGCTCGACACGCTGCAGAACACCGCAGCT 3641
RESULT 7
LAMCG 48502 bp DNA circular PHG 17-APR-2002
LOCUS Bacteriophage lambda, complete genome.
DEFINITION J02459 M17233 M24325 V00636 X00906
ACCESSION J02459.1 GI:215104
VERSION DNA-binding protein; circular; coat protein; complete genome;
KEYWORDS origin of replication; repressor; unidentified reading frame.
SOURCE Bacteriophage lambda
ORGANISM Bacteriophage lambda
REFERENCE 1 (bases 1 to 12)
AUTHORS Wu, R. and Taylor, E.
TITLE Nucleotide sequence analysis of DNA. II. Complete nucleotide
sequence of the cohesive ends of bacteriophage lambda DNA
J. Mol. Biol. 57 (3), 491-511 (1971)
```

- [illegible]

- [illegible]

lambda DNA
JOURNAL Nature 272 (5652), 410-414 (1978)
MEDLINE 78135462
PUBMED 264238
REFERENCE 25 (bases 38212 to 38362)
AUTHORS Rosenberg, M., Court, D., Shimatake, H., Brady, C. and Wulff, D.L.
TITLE The relationship between function and DNA sequence in an interchromosomal regulatory region in phage lambda
JOURNAL Nature 272 (5652), 414-423 (1978)
MEDLINE 78135463
PUBMED 634366
REFERENCE 26 (bases 37224 to 37940)
AUTHORS Sauer, R.T.
TITLE DNA sequence of the bacteriophage gamma CI gene
JOURNAL Nature 276 (5685), 301-302 (1978)
MEDLINE 79053284
PUBMED 714163
REFERENCE 27 (bases 38597 to 39688)
AUTHORS Scherer, G.
TITLE Nucleotide sequence of the O gene and of the origin of replication in bacteriophage lambda DNA
JOURNAL Nucleic Acids Res. 5 (9), 3141-3156 (1978)
MEDLINE 79033241
PUBMED 704348
REFERENCE 28 (bases 29711 to 29811; 31043 to 31058)
AUTHORS Davies, R.W., Schreier, P.H. and Buchel, D.E.
TITLE Determination of the endpoints of partial deletion mutants of the attachment site of bacteriophage lambda by DNA sequencing
JOURNAL Nucleic Acids Res. 5 (9), 3209-3218 (1978)
MEDLINE 79033246
PUBMED 704352
REFERENCE 29 (bases 21661 to 31129)
AUTHORS Hoes, R.H. and Landy, A.
TITLE Structure of the lambda att sites generated by int-dependent deletions
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 75 (11), 5437-5441 (1978)
MEDLINE 79074833
PUBMED 364480
REFERENCE 30 (bases 39453 to 38500)

Query Match 82.8%; Score 585.4; DB 7; Length 48502;

Best Local Similarity 99.8%; Pred. No. 2.6e-148;

Matches 586; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 45 ATCTAGTCGATCAGGATCATATCGTCGGTCTTTTTCGGCTCAGTCATCGCCCAAGC 104
DB 3055 ATCCAGCTGCATCAGGATCATATCGTCGGTCTTTTTCGGCTCAGTCATCGCCCAAGC 3114
QY 105 TGGCGCTATCTGGGCATCGGGGAGGAAGACCGCCGCTTTTCCCGCGAGGTTGAAGCG 164
DB 3115 TGGCGCTATCTGGGCATCGGGGAGGAAGACCGCGTGCCTTTTCCCGCGAGGTTGAAGCG 3174
QY 165 GCATGAAAGAGTTTCCCGAGGATGACTGCTGCTGATTCAGCTTGAGCGAAACCGCAGC 224
DB 3175 GCATGAAAGAGTTTCCCGAGGATGACTGCTGCTGATTCAGCTTGAGCGAAACCGCAGC 3234
QY 225 TTTTACCATGATGATTCGGGAAGGTGGCCATCGACGCTTTTAAACCGTGAACGTTCGTT 284
DB 3235 TTTTACCATGATGATTCGGGAAGGTGGCCATCGACGCTTTTAAACCGTGAACGTTCGTT 3294
QY 285 CAGGCACCTGGGATACCACTGCTCGCGCTTTTCCGACACAGTTCCGGATGGTCAGC 344
DB 3295 CAGGCACCTGGGATACCACTGCTCGCGCTTTTCCGACACAGTTCCGGATGGTCAGC 3354
QY 345 CCGAAGCGCATCAGCAACCCGAACAATACCGCGCAGACGCGGAACCTGCCGTGCTG 404
DB 3355 CCGAAGCGCATCAGCAACCCGAACAATACCGCGCAGACGCGGAACCTGCCGTGCTG 3414
QY 405 CAGATTAATACAGCGGTGGCGCTGGGATATTAGTCAGCAGGACCGGTATCTCTGC 464
DB 3415 CAGATTAATACAGCGGTGGCGCTGGGATATTAGTCAGCAGGACCGGTATCTCTGC 3474
QY 465 TGGATGCCGAGAAATGGACATGATGATACCCCGTGAGTTATCCCGCGGGCGCGCTCGTTC 524

Db 3475 TGGATGCCGAGAAATGGACATGATGATACCCCGTGAGTTATCCCGCGGGCGCGCTCGTTC 3534
QY 525 ATTCAGCTTTTGAACCCCGTGGAGACGGGCGAGACTCGCGGTGCAAAATGTTTACAGC 584
Db 3535 ATTCAGCTTTTGAACCCCGTGGAGACGGGCGAGACTCGCGGTGCAAAATGTTTACAGC 3594
QY 585 GTGATGGACGATGAAGATGCTCGACACGCTGCGAGAACACGACGCT 631
Db 3595 GTGATGGACGATGAAGATGCTCGACACGCTGCGAGAACACGACGCT 3641
RESULT 8
AF434923
LOCUS Expression vector pIB1-LUC, complete sequence.
DEFINITION Expression vector pIB1-LUC, complete sequence.
ACCESSION AF434923
VERSION AF434923.1 GI:16904152
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 5969)
AUTHORS Zieler, H. and Huynh, C.Q.
TITLE Intron-dependent stimulation of marker gene expression in cultured insect cells
JOURNAL Insect Mol. Biol. 11 (1), 87-95 (2002)
MEDLINE 21830665
PUBMED 11841506
REFERENCE 2 (bases 1 to 5969)
AUTHORS Zieler, H. and Huynh, C.Q.
TITLE Direct Submision
JOURNAL Submitted (17-OCT-2001) Medical Entomology Section, Laboratory of Parasitic Diseases, National Institutes of Health, 4 Center Drive MSC 0425, Bethesda, MD 20892-0425, USA
FEATURES
source
1..5969
/organism="Expression vector pIB1-LUC"
/mol_type="genomic DNA"
/db_xref="taxon:176965"
/note="insect-specific"
97..614
/note="Autographa californica nuclear polyhedrosis virus hr5 enhancer"
615..1201
/note="Autographa californica nuclear polyhedrosis virus immediate early promoter"
1328..2980
/codon_start=1
/product="firefly luciferase"
/protein_id="AAL30778.1"
/db_xref="GI:16904153"
/translation="MEDAKNIKKGAPFPVPLEDTAGBQLHKAMKRYALVPGTIAFTD AHIEVNTIYAEYFENSRLAEAMKGLNTHRIVVCSNSIQFMPVLGALFIVAV APANDIYNERELLSNMQPTVFVSKKLOKILNVOKLP IIOKIIIMDSKTDYQG FOSMTFTVSHLPDGFNEVDVPFSDRDKTIALIWNSSGSLGKVALPRTACVR FSHARDPIRGNOI I P D T A I L S V V P R H G F G M F T T L G Y L I C G R V L M V R F E B E L F A C V R L O D Y I Q A L L V P T L F S P F A K S T L I D K Y D L S N L H E I A S G A P L S K E V G S A V A K R H L P G R Q Y G U L T E T T S A I L I T P E G D K F G A G K V V F F E A K V D L D T G T L G V N R G E L C V R G P T M S G V Y N N P E A T N A L I D K D G W L H S G D I A Y W D E D E H F F I V D R L K S L I K Y G Q V A P A E L S I L I Q H E N I F D A G V A G L P D D D A G E L P A A V V L E H G K T M T E K I V D Y V A S Q V T T A K K L G G V V F V D E V P K L T G L D A R K I R E L I I K A K K G G K S K L"
3146..3251
/note="Heliothis virescens p63 chaperonin gene double poly-adenylation sequence"
4168..5028
/codon_start=1
/product="beta-lactamase"
/protein_id="AAL30779.1"
/db_xref="GI:16904154"
/translation="MSIQHFRVALIPFPFAFCFLPVFAHPETLVKVKDAEDQLGARVGY IELDLSNGKILDESFRPEKRFPMWSTFKVLLCGAVLSRIIDAGQEQLGRRIRHYISNDLIVE YSPVTEKHLTDGMTVRELCSAAITWSDNTAANLLLTITGGPKELTAFLHNMGDHVTFL


```
QY 345 CCGAAGCGCATCAGAACCCGAAACAATACCGGCGACAGCCGGAACCTCGCGTGGGTG 404
Db 4021 CCGAAGCGCATCAGAACCCGAAACAATACCGGCGACAGCCGGAACCTCGCGTGGGTG 4080
QY 405 CAGATTAAATGACAGCGGTGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCTTGGC 464
Db 4081 CAGATTAAATGACAGCGGTGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCTTGGC 4140
QY 465 TGGATGCGCGAGAAATGACATGATACCCCGTGAGTATACCGCGCGCGCGCTCGTTC 524
Db 4141 TGGATGCGCGAGAAATGACATGATACCCCGTGAGTATACCGCGCGCGCGCTCGTTC 4200
QY 525 ATTACAGCTTTTGAACCCGTTGGAGGACCGGACAGACTCGCGGTGCAATGTGTTTACAGC 584
Db 4201 ATTACAGCTTTTGAACCCGTTGGAGGACCGGACAGACTCGCGGTGCAATGTGTTTACAGC 4260
QY 585 GTGATGGAGCAGATGAAGATGCTCGACACCGCTCGAGAACAC 625
Db 4261 GTGATGGAGCAGATGAAGATGCTCGACACCGCTCGATGAAC 4301

RESULT 10
AF434926
LOCUS
DEFINITION Expression vector 409-MUT, DNA circular SYN 04-APR-2002
ACCESSION AF434926
VERSION AF434926.1 GI:16904161
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE Expression vector 409-MUT
Expression vector 409-MUT
other sequences; artificial sequences; vectors.
Zieler, H. and Huynh, C.Q.
intron-dependent stimulation of marker gene expression in cultured
insect cells
Insect Mol. Biol. 11 (1), 87-95 (2002)
MEDLINE 21830665
PUBMED 11841506
REFERENCE
AUTHORS Zieler, H. and Huynh, C.Q.
TITLE Direct Submission
JOURNAL Submitted (17-OCT-2001) Medical Entomology Section, Laboratory of
Parasitic Diseases, National Institutes of Health, 4 Center Drive
MSC 0425, Bethesda, MD 20892-0425, USA
LOCATION/Qualifiers
FEATURES
source
1. .6182
/organism="Expression vector 409-MUT"
/mol_type="genomic DNA"
/db_xref="taxon:176968"
/notes="insect-specific"
97. .614
/notes="Autographa californica nuclear polyhedrosis virus
hr5 enhancer"
enhancer
615. .1201
/notes="Autographa californica nuclear polyhedrosis virus
immediate early promoter"
promoter
1266. .1398
/notes="artificial intron with mutated 3' splice site"
intron
1541. .3193
/codon_start=1
/product="firefly luciferase"
/protein_id="AAL30784.1"
/db_xref="GI:16904162"
CDS
/translations="MEDAKNKKGPAPFPYLEDCTAGEQLHKMKRYALVPGTIATFD
AHLENIYTAAYFEMSVLAEAMKRYGLNTNHRIVVCSSENSLOFFMPVIGALFIGAV
APANDIYNRELNSMNIQSPTVFSVSKGLQKILNVOKKLPKIIIMDSKTDYQG
FSQMYFTVFNQIIPDPTAILSVVPHFGFMFTTLGLVLCGPRVLMYRFEELFLRS
LQDYKQSLALPTLSPFAKSLIDKVDLSNHEIASGAPLSKEVGAIVAKRFLHP
GIRQGGLETTSAILITPEGDDKPGAVGVFFFAKVDLDLTGKTLGVNQRELVCY
RGPIMSVGVNNEATNALIDKWLHSGDIAVDEHFFIVDRKLSLKYGYQVA
PAELSLILQHPNIFDAGVAGLPDPAVAVVLEHGKNTKEIIVDYVASQVTT
AKKLRGVWFVDPVKLTGKLDARKIREILIKAKKGGSKL"
```

```
misc_feature 3359. .3464
/notes="Heliothis virescens p63 chaperonin gene double
poly-adenylation sequence"
CDS
4381. .5241
/codon_start=1
/product="beta-lactamase"
/protein_id="AAL30785.1"
/db_xref="GI:16904163"
/translation="MSIQHPRVALIPFFPAFCIPVFAHPETLVKVKDAEDQLGARVGY
IEDLSNGKILSEFRPEERFPMWTFKVLGGAVLSRIDAGEQQLGRRIRHYQNDLVE
YSPVEKHLTDGMVRELCSAATMTSDNTAANLLLTITGGPKELTAFLPHMGDHVRL
DWEPELNEAIPIRDERDITTPVAMATTLKLTITGELLTLASQQLIDWMEADKVGAPL
LRSLALPAGFIADKSGAGERSGTIIAALGPDGKPSRIIVIVYITGSAQTMDSNRQIA
ETGASLIKHW"
5529. .5948
/notes="ColE1"
rep_origin
ORIGIN
Query Match 81.3%; Score 574.6; DB 12; Length 6182;
Best Local Similarity 99.3%; Pred. No. 1.8e-145;
Matches 577; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 45 ATCTAGCTCATCAGATCATATCGTCGGGTCTTTTTCGGCTCATGTCATCGCCCAAGC 104
Db 3721 ATCCAGCTCATCAGATCATATCGTCGGGTCTTTTTCGGCTCATGTCATCGCCCAAGC 3780
QY 105 TGGCGCTATCTGGGCATCGGGAGGAGAACCGCGTCCCTTTCCCGCAGGTTGAAGCG 164
Db 3781 TGGCGCTATCTGGGCATCGGGAGGAGAACCGCGTCCCTTTCCCGCAGGTTGAAGCG 3840
QY 165 GCATGAAAGAGTTTCCCGAGGATGACTCTCTGCTGCTATGAGCTTTGAGCGAAGAACG 224
Db 3841 GCATGAAAGAGTTTCCCGAGGATGACTCTCTGCTGCTGCTGAGCGTTCGAGCGAAG 3900
QY 225 TTACCATCATGATTCGGGAAGGTGTGCCATGCGACCTTTTAACGGTGAACCTGTTTCTG 284
Db 3901 TTACCATCATGATTCGGGAAGGTGTGCCATGCGACCTTTTAACGGTGAACCTGTTTCTG 3960
QY 285 CAGGCCACCTGGGATACACAGTTCTGTCGGGCTTTTCCGGACACAGTTCCGGATGTCAGC 344
Db 3961 CAGGCCACCTGGGATACACAGTTCTGTCGGGCTTTTCCGGACACAGTTCCGGATGTCAGC 4020
QY 345 CCGAAGCGCATCAGCAACCCGAAACAATACCGGCGACAGCCGGAACCTGCGGTGCGGTG 404
Db 4021 CCGAAGCGCATCAGCAACCCGAAACAATACCGGCGACAGCCGGAACCTGCGGTGCGGTG 4080
QY 405 CAGATTAAATGACAGCGGTGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCTCTGGC 464
Db 4081 CAGATTAAATGACAGCGGTGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCTCTGGC 4140
QY 465 TGGATGCGCGAGAAATGAGCATGATACCCCGTGAGTATACCGCGCGCGCGCTCGTTC 524
Db 4141 TGGATGCGCGAGAAATGAGCATGATACCCCGTGAGTATACCGCGCGCGCGCTCGTTC 4200
QY 525 ATTACAGCTTTTGAACCCCGTTGGAGGACCGGACAGACTCGCGGTGCAATGTGTTTACAGC 584
Db 4201 ATTACAGCTTTTGAACCCCGTTGGAGGACCGGACAGACTCGCGGTGCAATGTGTTTACAGC 4260
QY 585 GTGATGGAGCAGATGAAGATGCTCGACACCGCTCGAGAACAC 625
Db 4261 GTGATGGAGCAGATGAAGATGCTCGACACCGCTCGATGAAC 4301

RESULT 11
AF434927
LOCUS
DEFINITION Expression vector 410-FOR, complete sequence.
ACCESSION AF434927
VERSION AF434927.1 GI:16904164
KEYWORDS
SOURCE Expression vector 410-FOR
ORGANISM Expression vector 410-FOR
other sequences; artificial sequences; vectors.
```



```
REFERENCE 1 (bases 1 to 6182)
AUTHORS Zieler,H. and Huynh,C.Q.
TITLE Intron-dependent stimulation of marker gene expression in cultured
insect cells
JOURNAL Insect Mol. Biol. 11 (1), 87-95 (2002)
MEDLINE 21830665
PUBMED 11841506
REFERENCE 2 (bases 1 to 6182)
AUTHORS Zieler,H. and Huynh,C.Q.
TITLE Direct Submission
JOURNAL Submitted (17-OCT-2001) Medical Entomology Section, Laboratory of
Parasitic Diseases, National Institutes of Health, 4 Center Drive
MSC 0425, Bethesda, MD 20892-0425, USA
FEATURES
source
1. .6182
/organism="Expression vector 410-FOR"
/mol_type="genomic DNA"
/db_xref="taxon:176969"
/notes="insect-specific"
97. .614
enhancer
/notes="Autographa californica nuclear polyhedrosis virus
hr5 enhancer"
615. .1201
promoter
/notes="Autographa californica nuclear polyhedrosis virus
immediate early promoter"
1328. .2980
CDS
/codon_start=1
/product="firefly luciferase"
/protein_id="AAL30786.1"
/db_xref="GI:16904165"
/translaton="MEDAKNKKGPAPPYPLEDGTAGEQLHKAMKRYALVPGTIATFD
AHIEVNTIAYFEMSRLAEAMKRYGLNTHRIIVCSNSLQFPMVLGALFIGVAV
APANDIYRELNSMNIQSPTVVFVSKGLQKILNVOKLPIQKIIMDSKTDYQG
FQSMYTFVTHLPGFNEYDFPESFDKDTIALINSSGSTGLPKGVALPHRTACVR
PSHARDIPGNOI11PDAIILSWPHFGFMTTGLVLI CGFVLMVFEELFLRS
LDVYKIOSALVPTLFSFAKSTLI DKYDLSNHEIASGAPLSKEVGEAVAKRFLP
GIRQYGLTETTSAILITPEGDKPAGKVVPPFEAKVVDLDGTGLGVNQNGELCV
RPMIMSGYVNNPNTALIDKGLHSGDIAYWDEDEHFFI VDRLSKLIRKYGQVA
PAELSLILQHNPITFDAGVAGLPDDAGELPAAVYVLEHGKWTKEKIVDYVASQVTT
AKRLRGVWFVDEVPKGLTKLDARKIREILIKAKGKSKL"
3196. .3328
intron
/notes="artificial intron"
3359. .3464
misc_feature
/notes="Heliothis virescens p63 chaperonin gene double
poly-adenylation sequence"
4381. .5241
CDS
/codon_start=1
/product="beta-lactamase"
/protein_id="AAL30787.1"
/db_xref="GI:16904166"
/translaton="MSIQHFRVALIPFFAFLCFLVFAHPETLVKVKDAEDQLGARVGY
IELDLSGKILESFRPEFPMSTFKVLLCGAVLSRIDAGAEOLGRRIHYSNDLVE
YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTITIGPKELTAFLHNMGDVTRL
DRWEPELNEALPNDERTTPVAMATTLRLKLTGLLTSLASRQQLIDMWEADKVGGL
LRSLPAGWFIADKSGAGERGSRGIIAALGPDGKPSRIVVITYTGSQTMDERRRQIA
EIGASLIKHW"
5529. .5948
rep_origin
/notes="ColE1"
ORIGIN
Query Match 81.3%; Score 574.6; DB 12; Length 6182;
Best Local Similarity 99.3%; Pred. No. 1.8e-145;
Matches 577; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCCTTTTTCGGGCTCAGTCATCGCCCAAGC 104
DB 3721 ATCCAGCTGCATCAGGATCATATCGTCGGGTCCTTTTTCGGGCTCAGTCATCGCCCAAGC 3780
QY 105 TGGCGCTATCTGGGCATCGGGGAGGAAGAACCCCGTGCCTTTTCCCGCGAGGTTTGAAGCG 164
DB 3781 TGGCGCTATCTGGGCATCGGGGAGGAGAGACCCCGTGCCTTTTCCCGGAGGTTGAAGCG 3840
QY 165 GCATGGAAGAGTTTGGCGAGGATGACCTGCTGTCATGACGTGTGAGCGGAAAACGCACG 224
```

```
Db 3841 GCATGGAAGAGTTTGGCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 3900
QY 225 TTTACCATGATGATTTCGGGAAGGTGTGCCCATGCACGCTTTTAAACCGTGAACCTGTCGTT 284
DB 3901 TTTACCATGATGATTTCGGGAAGGTGTGCCCATGCACGCTTTTAAACCGTGAACCTGTCGTT 3960
QY 285 CAGGCCACTCTGGGATACCACTGTCGCGGCTTTTTCGGGACACAGTTCGGATCGTCAGC 344
DB 3961 CAGGCCACTCTGGGATACCACTGTCGCGGCTTTTTCGGGACACAGTTCGGATCGTCAGC 4020
QY 345 CCGAAGCGCATCAGCACACCCGAAACAATACCGGCGACACGCGGAACTCCGTCGCGGTGTG 404
DB 4021 CCGAAGCGCATCAGCACACCCGAAACAATACCGGCGACACGCGGAACTCCGTCGCGGTGTG 4080
QY 405 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTACGCGAGGACGGGTATCCTGGC 464
DB 4081 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTACGCGAGGACGGGTATCCTGGC 4140
QY 465 TGGATGCCGAGAAAATGGACATGATACCCCGTAGTATACCGCGGCGGCGGCTCGTTC 524
DB 4141 TGGATGCCGAGAAAATGGACATGATACCCCGTAGTATACCGCGGCGGCGGCTCGTTC 4200
QY 525 ATTACAGCTTTTGAACCCGTTGGAGGACGGGACGAGACTCGCGTGCAAATGTGTTTACAGC 584
DB 4201 ATTACAGCTTTTGAACCCGTTGGAGGACGGGACGAGACTCGCGTGCAAATGTGTTTACAGC 4260
QY 585 GTGATGAGCAGATGAAGATGCTCGACACGCTGCGAGAACAC 625
DB 4261 GTGATGAGCAGATGAAGATGCTCGACACGCTGCGAGAAC 4301
RESULT 12
AF434931
LOCUS AF434931 6182 bp DNA circular SYN 04-APR-2002
DEFINITION Expression vector 409-REV, complete sequence.
ACCESSION AF434931
VERSION AF434931.1 GI:16904176
KEYWORDS Expression vector 409-REV
SOURCE Expression vector 409-REV
ORGANISM Expression vector 409-REV
other sequences; artificial sequences; vectors.
REFERENCE 1 (bases 1 to 6182)
AUTHORS Zieler,H. and Huynh,C.Q.
TITLE Intron-dependent stimulation of marker gene expression in cultured
insect cells
JOURNAL Insect Mol. Biol. 11 (1), 87-95 (2002)
MEDLINE 21830665
PUBMED 11841506
REFERENCE 2 (bases 1 to 6182)
AUTHORS Zieler,H. and Huynh,C.Q.
TITLE Direct Submission
JOURNAL Submitted (17-OCT-2001) Medical Entomology Section, Laboratory of
Parasitic Diseases, National Institutes of Health, 4 Center Drive
MSC 0425, Bethesda, MD 20892-0425, USA
FEATURES
source
1. .6182
/organism="Expression vector 409-REV"
/mol_type="genomic DNA"
/db_xref="taxon:176973"
/notes="insect-specific"
97. .614
enhancer
/notes="Autographa californica nuclear polyhedrosis virus
hr5 enhancer"
615. .1201
promoter
/notes="Autographa californica nuclear polyhedrosis virus
immediate early promoter"
intron
complement(1236..1368)
/notes="artificial intron"
1541. .3193
CDS
/codon_start=1
/product="firefly luciferase"
/protein_id="AAL30794.1"
```

```
/db_xref="GI:16904177"
/translation="MEDAKNIKKGPAPFVPLEDTAGTEQLHKAMKRYALVPGTIAFTD
AHENITAEYFEMSVRLAEAMKRYGLNTHRIIVCSNSLOFPNVLGALFIVAV
APANDIYNERELLSNMISQPTVVFVSKKGLQILNVQKLP1IQKIIMDSKTDYQG
FSQMYTFVTSNQLTPPGNEVDYFVPSFDRDKTIALIMNSGSGTGLPKGVALPHRTACVR
FSHARDPIFGNOIIPDIALSVVPHHGFEMFTTGLYLCGFRVLMVYFEELFLURS
LODYKIQSALLVPTLFSFAKSLIDKYLNSLHEIASGAPLSKEVGEAVAKRHLPL
GIQOYGLTETTSAILITPEGDDKPGAVGVVFFPEAKVVDLDTGKTLGNORGEICV
RGPMSGVYNNPEATNALIDKDWLHSGDIAYWDEHFFIVDRKLSLKYKGQOVA
PAELSIILLOHPNIFDAGVAGLDDAGELPAAVVLEHGKMTTEKEIVDYVASQVTT
AKKRGVVVFVDEVPKGLTGKLDARKIREILLKAKKGGKSKL"
misc_feature
3359..3464
/note="Heliothis virescens p63 chaperonin gene double
poly-adenylation sequence"
CDS
4381..5241
/codon_start=1
/product="beta-lactamase"
/protein_id="AAL30795.1"
/db_xref="GI:16904178"
/translation="MSIOHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGY
IELDLNSGKILSFREPRFPMWSTFKVLLCGAVLSRIDAGQELGRRVHYSDNLDVE
YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLTTIGPKELTAFLHMGDHTVRL
DRWEPELNAIPNDERDTPMVAMATTLRKLLTGLLELLTLASRQQLIDWMEADKVAQPL
LRSLPAGMFIADKSGAGRGSRGIITAAIGDPGKPSRIIVYITGSGQATMDERNQRIA
EIGASLIKHW"
rep_origin
5529..5948
/note="Cole1"
ORIGIN
Query Match      81.3%; Score 574.6; DB 12; Length 6182;
Best Local Similarity 99.3%; Pred. No. 1.8e-145;
Matches 577; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 45 ATCTAGCTGATCAGGATCATATCGTCGGGTCTTTTTCGGGTCTAGTCATCGCCCAAGC 104
DB 3721 ATCCAGCTGATCAGGATCATATCGTCGGGTCTTTTTCGGGTCTAGTCATCGCCCAAGC 3780
QY 105 TGGCGCTATCTGGGCATCGGGAGGAAGACCGCGCTTTTCCCGCAGGTTTGAAGCG 164
DB 3781 TGGCGCTATCTGGGCATCGGGAGGAAGACCGCGCTTTTCCCGCAGGTTTGAAGCG 3840
QY 165 GCATGAAAGAGTTTCCGAGGATGACTGCTGCTGCAATTGACGTTGAGCGAAACGACG 224
DB 3841 GCATGAAAGAGTTTCCGAGGATGACTGCTGCTGCAATTGACGTTGAGCGAAACGACG 3900
QY 225 TTACCATGATGATTCGGGAAGGTGGCCATGCAAGCTTTTAAACGTTGAACGTTCGTT 284
DB 3901 TTACCATGATGATTCGGGAAGGTGGCCATGCAAGCTTTTAAACGTTGAACGTTCGTT 3960
QY 285 CAGGCCACCTGGGATACCACTGCTCGCGGCTTTTCCGACACAGATTCCGGATGGTCAGC 344
DB 3961 CAGGCCACCTGGGATACCACTGCTCGCGGCTTTTCCGACACAGATTCCGGATGGTCAGC 4020
QY 345 CCGAAGCGCATCAGCAACCGAACAATACCGGCGACAGCGGAACCTGCCGTGCCGGTGTG 404
DB 4021 CCGAAGCGCATCAGCAACCGAACAATACCGGCGACAGCGGAACCTGCCGTGCCGGTGTG 4080
QY 405 CAGATTAATGACAGCGTGGCGGCTGGGATATTAATGTCAGCAGAGCGGATCTCTGGC 464
DB 4081 CAGATTAATGACAGCGTGGCGGCTGGGATATTAATGTCAGCAGAGCGGATCTCTGGC 4140
QY 465 TGGATGCGCAGAAATGACATGATACCCCGTGAGTTACCGCGGGCGCGCTCGTTC 524
DB 4141 TGGATGCGCAGAAATGACATGATACCCCGTGAGTTACCGCGGGCGCGCTCGTTC 4200
QY 525 ATTACAGCTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAATGTGTTTACAGC 584
DB 4201 ATTACAGCTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAATGTGTTTACAGC 4260
QY 585 GTGATGAGCAGATGAAGATGCTCGACACCGCTCGAGACAC 625
DB 4261 GTGATGAGCAGATGAAGATGCTCGACACCGCTCGAGACAC 4301
```

```
RESULT 13
AF434932
LOCUS AF434932 6182 bp DNA circular SYN 04-APR-2002
DEFINITION Expression vector 410-REV, complete sequence.
ACCESSION AF434932
VERSION AF434932.1 GI:16904179
KEYWORDS
SOURCE Expression vector 410-REV
ORGANISM Expression vector 410-REV
OTHER SEQUENCES; artificial sequences; vectors.
REFERENCE 1 (bases 1 to 6182)
AUTHORS Zieler H. and Huyth, C.Q.
TITLE Intron-dependent stimulation of marker gene expression in cultured insect cells
JOURNAL Insect Mol. Biol. 11 (1), 87-95 (2002)
MEDLINE 21830665
PubMed 11841506
REFERENCE 2 (bases 1 to 6182)
AUTHORS Zieler H. and Huyth, C.Q.
TITLE Direct Submission
JOURNAL Submitted (17-OCT-2001) Medical Entomology Section, Laboratory of Parasitic Diseases, National Institutes of Health, 4 Center Drive MSC 0425, Bethesda, MD 20892-0425, USA
FEATURES
source
1..6182
/organism="Expression vector 410-REV"
/mol_type="genomic DNA"
/db_xref="taxon:176974"
/note="insect-specific"
97..614
/note="Autographa californica nuclear polyhedrosis virus hr-5 enhancer"
615..1201
/note="Autographa californica nuclear polyhedrosis virus immediate early promoter"
1328..2980
/codon_start=1
/product="firefly luciferase"
/protein_id="AAL30796.1"
/db_xref="GI:16904180"
/translation="MEDAKNIKKGPAPFVPLEDTAGTEQLHKAMKRYALVPGTIAFTD
AHENITAEYFEMSVRLAEAMKRYGLNTHRIIVCSNSLOFPNVLGALFIVAV
APANDIYNERELLSNMISQPTVVFVSKKGLQILNVQKLP1IQKIIMDSKTDYQG
FSQMYTFVTSNQLTPPGNEVDYFVPSFDRDKTIALIMNSGSGTGLPKGVALPHRTACVR
FSHARDPIFGNOIIPDIALSVVPHHGFEMFTTGLYLCGFRVLMVYFEELFLURS
LODYKIQSALLVPTLFSFAKSLIDKYLNSLHEIASGAPLSKEVGEAVAKRHLPL
GIQOYGLTETTSAILITPEGDDKPGAVGVVFFPEAKVVDLDTGKTLGNORGEICV
RGPMSGVYNNPEATNALIDKDWLHSGDIAYWDEHFFIVDRKLSLKYKGQOVA
PAELSIILLOHPNIFDAGVAGLDDAGELPAAVVLEHGKMTTEKEIVDYVASQVTT
AKKRGVVVFVDEVPKGLTGKLDARKIREILLKAKKGGKSKL"
complement (3166..3298)
/note="artificial intron"
misc_feature
3359..3464
/note="Heliothis virescens p63 chaperonin gene double
poly-adenylation sequence"
4381..5241
/codon_start=1
/product="beta-lactamase"
/protein_id="AAL30797.1"
/db_xref="GI:16904181"
/translation="MSIOHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGY
IELDLNSGKILSFREPRFPMWSTFKVLLCGAVLSRIDAGQELGRRVHYSDNLDVE
YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLTTIGPKELTAFLHMGDHTVRL
DRWEPELNAIPNDERDTPMVAMATTLRKLLTGLLELLTLASRQQLIDWMEADKVAQPL
LRSLPAGMFIADKSGAGRGSRGIITAAIGDPGKPSRIIVYITGSGQATMDERNQRIA
EIGASLIKHW"
5529..5948
/note="Cole1"
rep_origin
81.3%; Score 574.6; DB 12; Length 6182;
Query Match
Best Local Similarity 99.3%; Pred. No. 1.8e-145;
```


Matches		577;	Conservative	0;	Mismatches	4;	Indels	0;	Gaps	0;
Qy	45	ATCTAGCTGCATCAGGATCATATCGTCGGGCTCTTTTTCGGGCTCAGTCATCGCCCAAGC	104							
Db	3721	ATCCAGCTGCATCAGGATCATATCGTCGGGCTCTTTTTCGGGCTCAGTCATCGCCCAAGC	3780							
Qy	105	TGGCGCTATCTGGGCATCGGGGAGGAGAGCGCGTGCCTTTTTCGGGAGGTTGAAGCG	164							
Db	3781	TGGCGCTATCTGGGCATCGGGGAGGAGAGCGCGTGCCTTTTTCGGGAGGTTGAAGCG	3840							
Qy	165	GCATGGAAAGAGTTTCCGAGGATGACTGCTGCTGCTTGCATTGACGTTGAGCGAAAAACGACG	224							
Db	3841	GCATGGAAAGAGTTTCCGAGGATGACTGCTGCTGCTTGCATTGACGTTGAGCGAAAAACGACG	3900							
Qy	225	TTTACCATGATGATTCCGGAAGAGTGTGCCATGCACGCTTTTAAACGTTGAACTGTTCTGTT	284							
Db	3901	TTTACCATGATGATTCCGGAAGAGTGTGCCATGCACGCTTTTAAACGTTGAACTGTTCTGTT	3960							
Qy	285	CAGGCCACCTGGGATACCACTGCTCGGGCTTTTTCGGACACAGTTCCGGATGGTCAGC	344							
Db	3961	CAGGCCACCTGGGATACCACTGCTCGGGCTTTTTCGGACACAGTTCCGGATGGTCAGC	4020							
Qy	345	CCGAAGCGCATCAGCAACCCGAAACAATACCGCGACAGCGGAACTGCGCGGCTGCGGCTG	404							
Db	4021	CCGAAGCGCATCAGCAACCCGAAACAATACCGCGACAGCGGAACTGCGCGGCTGCGGCTG	4080							
Qy	405	CAGATTAATGACAGCGGTGCGCGCTGGGATATTACGTCACGAGGACGGGTATCTCTGGC	464							
Db	4081	CAGATTAATGACAGCGGTGCGCGCTGGGATATTACGTCACGAGGACGGGTATCTCTGGC	4140							
Qy	465	TGGATGCGCGAGAAATGACATGATACCCGTCGAGTTACCGCGGGCGCGCTCGTTTC	524							
Db	4141	TGGATGCGCGAGAAATGACATGATACCCGTCGAGTTACCGCGGGCGCGCTCGTTTC	4200							
Qy	525	ATTACAGCTTTTGAACCGTGGAGAGCGGCGAGACTCGCGTGCAAAATGTTTACAGC	584							
Db	4201	ATTACAGCTTTTGAACCGTGGAGAGCGGCGAGACTCGCGTGCAAAATGTTTACAGC	4260							
Qy	585	GTGATGAGCAGATGAAGATGCTGCACACGCTGCAGAACAC	625							
Db	4261	GTGATGAGCAGATGAAGATGCTGCACACGCTGCATGAAC	4301							
RESULT 14										
AF434924										
LOCUS										
DEFINITION Expression vector pACTIN-LUC, complete sequence.										
ACCESSION AF434924										
VERSION AF434924.1 GI:16904155										
KEYWORDS										
SOURCE										
ORGANISM										
REFERENCE										
AUTHORS Zieler, H. and Huynh, C.Q.										
TITLE Intron-dependent stimulation of marker gene expression in cultured insect cells										
JOURNAL Insect Mol. Biol. 11 (1), 87-95 (2002)										
MEDLINE 21830565										
PUBMED 11841506										
REFERENCE										
AUTHORS Zieler, H. and Huynh, C.Q.										
TITLE Direct Submission										
JOURNAL Submitted (17-OCT-2001) Medical Entomology Section, Laboratory of Parasitic Diseases, National Institutes of Health, 4 Center Drive MSC 0425, Bethesda, MD 20892-0425, USA										
FEATURES										
source										
1. 7421										
/organism="Expression vector pACTIN-LUC"										
/mol_type="genomic DNA"										
/db_xref="taxon:176966"										
/note="insect-specific"										
103. 2660										
promoter										

CDS		/note="Drosophila melanogaster actin5C promoter"	
2780. .4432		/codon_start=1	
/product="firefly luciferase"		/protein_id="AAL30780.1"	
/db_xref="GI:16904156"		/translation="MEDAKNIKKGPAPPFYPLEDGTAGEQLHKAKRVALVPGTIAFTD	
AHLEVNITTAAYEFEMSVRLAEAKWKYGLNTNHRIVVCSNSLQFMPVLGALFIGAV		APANDYINERELLNSNIQSPTVFSVSKGLKILNVQKLPITIKIIMDSKTDYQG	
FQMYTFVTSHPGPNFVDVPEFSDRDKTIALIMNSGSGTLPKGVLPKHTACVVR		FSHARDIPFGNQIIPDTAILSVVPFHFGFMFTTLGYLICGFRVLVMEBELFLPS	
LDYKIOSALLVPTLPSFPAKSLIDKYDLSNLHEIASGAPLSKEVGEAVAKRFLRP		GIRQGGLETTSAILITPEGDDKPGAVKVPFPFPAKVVDLDTGKTLGVNQGELCV	
RGPMISGVYNNPEATLIDKDWLHSGDIAWDEDEHFFIVDRLSKLLIKYGYQVA		PAELSESILQHPNIFDAGVAGLPDDAGELPAAVVLVLEHGTMTKEBIVDYNASQVTT	
AKKLGGVVFVDVPPKGLTGKLDARKIREILLIKAKKGKSKL"		4598. .4703	
misc_feature		/note="Heliothis virescens p63 chaperonin gene double	
poly-adenylation sequence"		5620. .6480	
CDS		/codon_start=1	
/product="beta-lactamase"		/protein_id="AAL30781.1"	
/db_xref="GI:16904157"		/translation="MSIQHFRVALIPFFAAFCPLVPFAHPETLVKVKDAEDQLGARVGY	
IELDLSNGKILLESFRPEERFPMSTDKTAVLSRIDAGQQLGRIIRHISQNDLVR		YSPVTEKHITLDTGKMTRELCSAAITMSDNTAANLLITIGGKELTAFHNMGDHVLRL	
DWPELNEAI PNDRDITMPVAMATTLRKLTGELLTLASRQQLIDMDEADKVGPL		LRSLAPAGWFIADKSGAGERSGRIIAALGPDGKPSRIVIYITGSGATWDERNRQIA	
EIGASLIKKHW"		6788. .7187	
rep_origin		/note="ColE1"	
ORIGIN		Query Match	
Best Local Similarity		81.3%; Score 574.6; DB 12; Length 7421;	
Matches		577; Conservative 0; Mismatches 4; Indels 0; Gaps 0;	
Qy	45	ATCTAGCTGCATCAGGATCATATCGTCGGGCTCTTTTTCGGGCTCAGTCATCGCCCAAGC	104
Db	4960	ATCCAGCTGCATCAGGATCATATCGTCGGGCTCTTTTTCGGGCTCAGTCATCGCCCAAGC	5019
Qy	105	TGGCGCTATCTGGGCATCGGGGAGGAGAACCCGTCGCTTTTCCCGCGAGGTTGAAGCG	164
Db	5020	TGGCGCTATCTGGGCATCGGGGAGGAGAACCCGTCGCTTTTCCCGCGAGGTTGAAGCG	5079
Qy	165	GCATGGAAAGAGTTTCCGAGGATGACTGCTGCTGCAATTCACGTTGAGCGAAAAACGACG	224
Db	5080	GCATGGAAAGAGTTTCCGAGGATGACTGCTGCTGCAATTCACGTTGAGCGAAAAACGACG	5139
Qy	225	TTTACCATGATGATTCGGGAAGGTGTGGCCATGCGACGCTTTTAAACGTTGAACTGTTCTGTT	284
Db	5140	TTTACCATGATGATTCGGGAAGGTGTGGCCATGCGACGCTTTTAAACGTTGAACTGTTCTGTT	5199
Qy	285	CAGGCCACCTGGGATACCACTGCTCGGGCTTTTTCGGACACAGTTCCGGATGGTCAGC	344
Db	5200	CAGGCCACCTGGGATACCACTGCTCGGGCTTTTTCGGACACAGTTCCGGATGGTCAGC	5259
Qy	345	CCGAAGCGCATCAGCAACCCGAAACAATACCGCGACAGCGGAACTGCGCGGCTGCGGCTG	404
Db	5260	CCGAAGCGCATCAGCAACCCGAAACAATACCGCGACAGCGGAACTGCGCGGCTGCGGCTG	5319
Qy	405	CAGATTAATGACAGCGGTGCGCGCTGGGATATTACGTCACGAGGACGGGTATCTCTGGC	464
Db	5320	CAGATTAATGACAGCGGTGCGCGCTGGGATATTACGTCACGAGGACGGGTATCTCTGGC	5379
Qy	465	TGGATGCGCGAGAAATGACATGATACCCGTCGAGTTACCGCGGGCGCGCTCGTTTC	524
Db	5380	TGGATGCGCGAGAAATGACATGATACCCGTCGAGTTACCGCGGGCGCGCTCGTTTC	5439
Qy	525	ATTACAGCTTTTGAACCCGTCGAGGAGCGGCGAGACTCGCGTGCAAAATGTTTACAGC	584
Db	5440	ATTACAGCTTTTGAACCCGTCGAGGAGCGGCGAGACTCGCGTGCAAAATGTTTACAGC	5499

[illegible]

Search completed: June 5, 2005, 19:22:01
Job time : 2967.77 secs

QY	585	GTGATGAGCAGATGAAGATGCTCGACACGCTCGAGAACAC	625
Db	5500	GTGATGAGCAGATGAAGATGCTCGACACGCTCGATGAAC	5540
RESULT 15			
AF434928			
LOCUS	AF434928	7634 bp DNA	circular SYN 04-APR-2002
DEFINITION	Expression vector 411-FOR, complete sequence.		
ACCESSION	AF434928		
VERSION	AF434928.1	GI:16904167	
KEYWORDS			
SOURCE	Expression vector 411-FOR		
ORGANISM	Expression vector 411-FOR		
REFERENCE	Other sequences; artificial sequences; vectors.		
AUTHORS	1 (bases 1 to 7634)		
TITLE	Zieler, H. and Huynh, C.Q.		
JOURNAL	Intron-dependent stimulation of marker gene expression in cultured insect cells		
MEDLINE	Insect Mol. Biol. 11 (1), 87-95 (2002)		
PUBMED	21830665		
REFERENCE	11841506		
AUTHORS	2 (bases 1 to 7634)		
TITLE	Zieler, H. and Huynh, C.Q.		
JOURNAL	Submitted (17-OCT-2001) Medical Entomology Section, Laboratory of Parasitic Diseases, National Institutes of Health, 4 Center Drive MSC 0425, Bethesda, MD 20892-0425, USA		
FEATURES	Location/Qualifiers		
source	1..7634		
	/organism="Expression vector 411-FOR"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:176970"		
	/note="insect-specific"		
promoter	103..2660		
	/note="Drosophila melanogaster actin5C promoter"		
intron	2718..2850		
	/note="artificial intron"		
CDS	2993..4645		
	/codon_start=1		
	/product="firefly luciferase"		
	/protein_id="AAL30788.1"		
	/db_xref="GI:16904168"		
	/translation="MEDAKNIKGPAPFVPLEDCTAGEQLHKMKRYALVPGTIATFD AHLEVALTYAYPEMSVRKAEKRYGLNTNHRIVVCSNSLQFMPVLGALFVIGVAF APANDYINRELNSNNI SPTVYFVSKGLQILNVKQKPIIQKIIIMDSKTDYQV FQSMYFTVSLPPGNEVD FVPSFDRDKTALIMNSGSLGPKGVALPHRTACV FSHARDIPFQNIIPITAILSVVPFHFGMFTTGLYICGFRVLWYRFEELFVR LQDYKQSALLVPTLSPFAKSTLIDKYDLSNLHEIASGAPLSKEVGAVALRFLP GIRQYGLTETTSAILITPEGDKPGAVKVPFFKAVVDLDGTGLGVNORGELOCV RGMMSGYVNNFEATNALIDKQWLHSGDIAYWDEDEHFFIVDRLSKLIKYGYQVA PALESILQHPNIDFAGVAGLDPDAGEDPAAVVLEHGKMTMTEKEIVDIYVASQVIT AKKRGGVFVDPVPGKLTGLDKARKIREILIKAKKGKSKL"		
misc_feature	4811..4916		
	/note="Heliothis virescens p63 chaperonin gene double poly-adenylation sequence"		
CDS	5833..6693		
	/codon_start=1		
	/product="beta-lactamase"		
	/protein_id="AAL30789.1"		
	/db_xref="GI:16904169"		
	/translation="MSGTQHFVRVALIPFFAFCFLPVFAHPETLVKVKDAEDQLGARVY IEIDLNSGKILSFPRPEERPMWSFKVLICGVALRIDAGQOLGRRIHYSONDLVE YSPVTEKHLDGMGTVELCSAAITMSDNTAANLLTTIGCKELTAFLNMGHPVTRL DRWPELNEAIPNDERDITMPVAMATLRLKLLTGLSRLQLIDNMEDAKVAGPL LNSALPAGFWFIADKSGSGEKGRGIIAALGFDGKPSRIVVIYITGSGATMDERNRQIA EIGASLIKHW"		
rep_origin	6981..7400		
	/note="ColE1"		
ORIGIN			

Query Match 81.3%; Score 574.6; DB 12; Length 7634;

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 5, 2005, 13:39:19 ; Search time 117.806 Seconds
(without alignments)
9819.908 Million cell updates/sec

Title: US-10-826-523-40
Perfect score: 707
Sequence: 1 ggaatccatgcgtcaatttt.....aaaattgacgcattggatcc 707

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTCTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	566.2	80.1	46819	3	US-09-453-702B-72
C 2	561.4	79.4	22306	3	US-09-453-702B-251
C 3	223	31.5	34083	3	US-09-453-702B-96
C 4	73.8	10.4	5194	3	US-08-844-274-16
C 5	73.8	10.4	5194	4	US-09-598-421-16
C 6	73.4	10.4	5194	3	US-08-844-274-17
C 7	73.4	10.4	5194	4	US-09-598-421-17
8	67.8	9.6	5679	3	US-08-844-274-10
9	67.8	9.6	5679	4	US-09-598-421-10
10	67.8	9.6	6723	3	US-08-844-274-13
C 11	67.8	9.6	6723	3	US-08-844-274-14
C 12	67.8	9.6	6723	4	US-09-598-421-13
C 13	67.8	9.6	6723	4	US-09-598-421-14
C 14	67.8	9.6	7560	3	US-08-844-274-20
C 15	67.8	9.6	7560	4	US-09-598-421-20
C 16	67.8	9.6	9423	4	US-09-377-066-6
C 17	64.4	9.1	6448	3	US-08-844-274-15
C 18	64.4	9.1	6448	4	US-09-598-421-15
C 19	63.8	9.0	2476	3	US-08-844-274-11
C 20	63.8	9.0	2476	4	US-09-598-421-11
21	39.8	5.6	489	4	US-09-252-991A-15689
22	39.8	5.6	888	4	US-09-252-991A-15722
C 23	39.8	5.6	2001	4	US-09-252-991A-15722
C 24	39.8	5.6	2475	4	US-09-252-991A-15589
C 25	39.2	5.5	1734	6	US-09-252-991A-15758
C 26	39.2	5.5	1734	6	5352575-8
C 27	35.4	5.0	209210	4	5352575-8
28					US-09-949-016-15094

Sequence 3154, Ap
Sequence 2816, Ap
Sequence 2, Appl
Sequence 1, Appl
Sequence 7848, Ap
Sequence 1011, Ap
Sequence 1011, Ap
Sequence 1011, Ap
Sequence 1011, Ap
Sequence 1011, Ap
Sequence 1011, Ap
Sequence 785, App
Sequence 1230, Ap
Sequence 573, App
Sequence 1868, Ap
Sequence 17150, A
Sequence 2, Appl
Sequence 2, Appl
Sequence 2, Appl

33.8 4.8 1254 4 US-09-252-991A-3154
33.8 4.8 1692 4 US-09-252-991A-2816
33.2 4.7 4403765 3 US-09-103-840A-2
33.2 4.7 4411529 3 US-09-103-840A-1
33.2 4.7 408 4 US-09-902-540-7848
33.2 4.7 449 4 US-09-702-705-1011
33.2 4.7 449 4 US-09-736-457-1011
33.2 4.7 449 4 US-09-614-124B-1011
33.2 4.7 449 4 US-09-671-325-1011
33.2 4.7 449 4 US-09-658-824-1011
33.2 4.7 449 4 US-09-902-540-785
33.2 4.7 5125 4 US-09-902-540-1230
33.2 4.7 24754 4 US-09-799-451-573
33.2 4.6 1689 4 US-09-270-767-1868
32.8 4.6 2823 4 US-09-270-767-17150
32.8 4.6 3468 1 US-07-951-715A-2
32.8 4.6 3468 2 US-08-459-448A-2
32.8 4.6 3468 3 US-08-459-595A-2

ALIGNMENTS

RESULT 1
US-09-453-702B-72
; Sequence 72, Application US/09453702B
; Patent No. 6365723
; GENERAL INFORMATION:
; APPLICANT: Blattner, Frederick R.
; Burland, Nicole T.
; Perna, Nicole T.
; Plunkett, Guy
; Welch, Rod
; TITLE OF INVENTION: No. 6365723el Sequences of E. coli O157
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch. 1.44Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 8.0
CURRENT APPLICATION DATA: US/09/453,702B
APPLICATION NUMBER: US/09/453,702B
FILING DATE: 03-Dec-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/110,955
FILING DATE: 04-DEC-1998
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J.
REGISTRATION NUMBER: 27386
REFERENCE/DOCKET NUMBER: 960296.95017
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 251-5000
TELEFAX: (608) 251-9166
INFORMATION FOR SEQ ID NO: 72:
SEQUENCE CHARACTERISTICS:
LENGTH: 46819
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 72:
US-09-453-702B-72

Query Match 80.1%; Score 566.2; DB 3; Length 46819;
Best Local Similarity 97.8%; Pred. No. 6.6e-176;

Matches 574; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 45 ATCTAGCTCATCAGGATCATATCGTCGGGTCTTTTTCGGGCTCAGTCATCGCCCAAGC 104
|||
Db 14795 ATCCAGTCGATCAGGATCATATCGTCGGGTCTTTTTCGGGCTCAGTCATCGCCCAAGC 14854
|||
QY 105 TGGCGCTATCTGGGCTATCGGGGAGGAAGACCGTGCCTTTTCCCGCGAGGTTGAAGCG 164
|||
Db 14855 TGGCGCTATCTGGGCTATCGGGGAGGAAGACCGTGCCTTTTCCCGCGAGGTTGAAGCG 14914
|||
QY 165 GCATGAAAGAGTTTCGCGAGGATGACTGTCTGTCATTGACGTTGAGCGAAGAACGACG 224
|||
Db 14915 GCATGAAAGAGTTTCGCGAGGATGACTGTCTGTCATTGACGTTGAGCGAAGAACGACG 14974
|||
QY 225 TTTACCATGATGATTTCGGGAGGTGTGGCCATGCACGCTTTAAACGTTGAACTGTTT 284
|||
Db 14975 TTTACCATGATGATTTCGGGAGGTGTGGCCATGCACGCTTTAAACGTTGAACTGTTT 15034
|||
QY 285 CAGGCCACCTGGGATACCAAGTTTCGTCGGCTTTTCCGGACACAGTTCCGGATGTCAGC 344
|||
Db 15035 CAGGCCACCTGGGATACCAAGTTTCGTCGGCTTTTCCGGACACAGTTCCGGATGTCAGC 15094
|||
QY 345 CCGAAGCGCATCAGCAACCGGAACAATACCGGCGACAGCGGAACTGCCGTGCCGTGTG 404
|||
Db 15095 CCGAAGCGCATCAGCAACCGGAACAATACCGGCGACAGCGGAACTGCCGTGCCGTGTG 15154
|||
QY 405 CAGATTAAATGACAGCGGTGGCGCTGGGATATTAGCTCAGCGAGACGGGTATCTGGC 464
|||
Db 15155 CAGATTAAATGACAGCGGTGGCGCTGGGATATTAGCTCAGCGAGACGGGTATCTGGC 15214
|||
QY 465 TGGATGCCGAGAAATGGACATGATACCCCGTAGTTTACCGCGCGCGCTCGTTTC 524
|||
Db 15215 TGGATGCCGAGAAATGGACATGATACCCCGTAGTTTACCGCGCGCGCTCGTTTC 15274
|||
QY 525 ATTACGTTTTTGAAACCGGTGGAGGACGGGACAGACTCGCGGTGCAAAATGTGTTTACAGC 584
|||
Db 15275 ATTACGTTTTTGAAACCGGTGGAGGACGGGACAGACTCGCGGTGCAAAATGTGTTTACAGC 15334
|||
QY 585 GTGATGGACGATGAAGATGCTCGACACGCTGCAGAACACCGCAGCT 631
|||
Db 15335 GTGATGGACGATGAAGATGCTCGACACGCTGCAGAACACCGCAGCT 15381
|||

RESULT 2

US-09-453-702B-251/c

; Sequence 251, Application US/09453702B

; Patent No. 6365723

; GENERAL INFORMATION:

; APPLICANT: Blattner, Frederick R.

; Burland, Valerie

; Perna, Nicole T.

; Plunkett, Guy

; Welch, Rod

; TITLE OF INVENTION: No. 6365723el Sequences of E. coli O157

; NUMBER OF SEQUENCES: 265

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Quarles & Brady

; STREET: 1 South Pinckney Street

; CITY: Madison

; STATE: WI

; COUNTRY: US

; ZIP: 53701-2113

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch. 1.44Mb storage

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Word Perfect 8.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/453,702B

; FILING DATE: 03-Dec-1999

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 60/110,955

; FILING DATE: 04-DEC-1998

; ATTORNEY/AGENT INFORMATION:

; NAME: Seay, Nicholas J.

; REGISTRATION NUMBER: 27386

; REFERENCE/DOCKET NUMBER: 960296.95017

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (608) 251-5000

; TELEFAX: (608) 251-9166

; INFORMATION FOR SEQ ID NO: 251:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 22306

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; SEQUENCE DESCRIPTION: SEQ ID NO: 251:

US-09-453-702B-251

Query Match 79.4%; Score 561.4; DB 3; Length 22306;

Best Local Similarity 97.3%; Pred. No. 1.7e-174;

Matches 571; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 45 ATCTAGCTCATCAGGATCATATCGTCGGGTCTTTTTCGGGCTCAGTCATCGCCCAAGC 104
|||
Db 17193 ATCCAGTTGCATCAGGATCATATCGTCGGGTCTTTTTCGGGCTCAGTCATCGCCCAAGC 17134
|||
QY 105 TGGCGCTATCTGGGCTATCGGGGAGGAAGACCGTGCCTTTTCCCGCGAGGTTGAAGCG 164
|||
Db 17133 TGGCGCTATCTGGGCTATCGGGGAGGAAGACCGTGCCTTTTCCCGCGAGGTTGAAGCG 17074
|||
QY 165 GCATGAAAGAGTTTGGCCGAGGATGACTGCTGCTCATTTGAGCGTGTAGCGAAGAACGACG 224
|||
Db 17073 GCATGAAAGAGTTTGGCCGAGGATGACTGCTGCTCATTTGAGCGTGTAGCGAAGAACGACG 17014
|||
QY 225 TTTACCATGATGATTTCGGGAAAGGTGTGGCCATGCACGCTTTTAAACGTTGAACTGTTT 284
|||
Db 17013 TTTACCATGATGATTTCGGGAAAGGTGTGGCCATGCACGCTTTTAAACGTTGAACTGTTT 16954
|||
QY 285 CAGGCCACCTGGGATACCAAGTTTCGTCGGCTTTTCCGGACACAGTTCCGGATGTCAGC 344
|||
Db 16953 CAGGCCACCTGGGATACCAAGTTTCGTCGGCTTTTCCGGACACAGTTCCGGATGTCAGC 16894
|||
QY 345 CCGAAGCGCATCAGCAACCGGAACAATACCGGCGACAGCGGAACTGCCGTGCCGTGTG 404
|||
Db 16893 CCGAAGCGCATCAGCAACCGGAACAATACCGGCGACAGCGGAACTGCCGTGCCGTGTG 16834
|||
QY 405 CAGATTAAATGACAGCGGTGGCGCTGGGATATTAGCTCAGCGAGACGGGTATCTGGC 464
|||
Db 16833 CAGATTAAATGACAGCGGTGGCGCTGGGATATTAGCTCAGCGAGACGGGTATCTGGC 16774
|||
QY 465 TGGATGCCGAGAAATGGACATGATACCCCGTAGTTTACCGCGCGCGCTCGTTTC 524
|||
Db 16773 TGGATGCCGAGAAATGGACATGATACCCCGTAGTTTACCGCGCGCGCTCGTTTC 16714
|||
QY 525 ATTACAGTTTTTGAACCGGTGGAGGACGGGACAGACTCGCGGTGCAAAATGTGTTTACAGC 584
|||
Db 16713 ATTACAGTTTTTGAACCGGTGGAGGACGGGACAGACTCGCGGTGCAAAATGTGTTTACAGC 16654
|||
QY 585 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACCGCAGCT 631
|||
Db 16653 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACCGCAGCT 16607
|||

RESULT 3

US-09-453-702B-96/c

; Sequence 96, Application US/09453702B

; Patent No. 6365723

; GENERAL INFORMATION:

; APPLICANT: Blattner, Frederick R.

; Burland, Valerie

; Perna, Nicole T.

; Plunkett, Guy

; Welch, Rod

TITLE OF INVENTION: No. 6365723el Sequences of E. coli Q157
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESSEE: Quarles & Brady
STREET: 1 South Pinckney Street
CITY: Madison
STATE: WI
COUNTRY: US
ZIP: 53701-2113
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch. 1.44Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 8.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/453,702B
FILING DATE: 03-Dec-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/110,955
FILING DATE: 04-DEC-1998
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J.
REGISTRATION NUMBER: 27386
REFERENCE/DOCKET NUMBER: 960296.95017
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 251-5000
TELEFAX: (608) 251-9166
INFORMATION FOR SEQ ID NO: 96:
SEQUENCE CHARACTERISTICS:
LENGTH: 34063
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 96:
US-09-453-702B-96

```

Qy 531 GTTTTGAACCCGTGAGAGACGGGACAGCTCCGGTGCAAATGCTTTTACAGCGTGATG 590
Db 14215 ATTTTTCGACCCGTGAGAGACGGGACAGACCCCGTGGGGCCAAACCAGTTTTTACAGCGTCA TG 14155
Qy 591 GACACAGATGAAGATGCTGCACACGCTGCAGAAACACGCAGCTAGATTAACCC 641
Db 14155 GRAACGCTGAAGATGCTCGATTCCCTGCAGGCAACACAGCTTCAGTCGGCC 14105

RESULT 4
US-08-844-274-16/c
; Sequence 16, Application US/08844274B
; Patent No. 6218185
; GENERAL INFORMATION:
; APPLICANT: Fraser Jr., Malcom J.
; APPLICANT: Shirk, Paul D.
; APPLICANT: Elick, Teri A.
; APPLICANT: Perera, Omachthage
; TITLE OF INVENTION: PiggyBac Transposon-Based Genetic Transformation System
; TITLE OF INVENTION: for Insects
; FILE REFERENCE: 0148.96
; CURRENT APPLICATION NUMBER: US/08/844,274B
; CURRENT FILING DATE: 1997-04-18
; EARLIER APPLICATION NUMBER: 60/016,234
; EARLIER FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO 16
; LENGTH: 5194
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:plasmid
; OTHER INFORMATION: IFF2B/Xpuc18.1
US-08-844-274-16

```

```

RESULT 5
US-09-598-421-16/c
; Sequence 16. Application US/09598421
; Patent No. 6551825
; GENERAL INFORMATION:
; APPLICANT: Fraser Jr., Malcom J.
; APPLICANT: Shirk, Paul D.
; APPLICANT: Elick, Teri A.
; APPLICANT: Perera, Omaththage
; TITLE OF INVENTION: PiggyBac Transposon-Based Genetic Transformation System
; FILE REFERENCE: 0148.96
; CURRENT APPLICATION NUMBER: US/09/598,421
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: 60/016,234
; PRIOR FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO 16
; LENGTH: 5194
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:

```

```
; OTHER INFORMATION: Description of Artificial Sequence:plasmid
; OTHER INFORMATION: IFF2B/Xpuc18.1
US-09-598-421-16

Query Match      10.4%; Score 73.8; DB 4; Length 5194;
Best Local Similarity 83.2%; Pred. No. 2.3e-13;
Matches 84; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 605 GCTCGACACGCTCGAGAACACGCGAGCTAGATTAAACCTAGAAAGATAATCATATTGTGAC 664
Db 2955 GCTCGGTACCCGGGGATCGGATCTCTAGATTAAACCTAGAAAGATAATCATATTGTGAC 2896

QY 665 GTAGCTTAAAGATAATCATCGTAAATAATTGACGCATGGGAT 705
Db 2895 GTAGCTTAAAGATAATCATCGTAAATAATTGACGCATGTGTT 2855

RESULT 6
US-08-844-274-17/c
; Sequence 17, Application US/08844274B
; Patent No. 6218185
; GENERAL INFORMATION:
; APPLICANT: Fraser Jr., Malcom J.
; APPLICANT: Shirk, Paul D.
; APPLICANT: Elick, Teri A.
; APPLICANT: Perera, Omaththage
; TITLE OF INVENTION: PiggyBac Transposon-Based Genetic Transformation System
; TITLE OF INVENTION: for Insects
; FILE REFERENCE: 0148.96
; CURRENT APPLICATION NUMBER: US/08/844,274B
; CURRENT FILING DATE: 1997-04-18
; EARLIER APPLICATION NUMBER: 60/016,234
; EARLIER FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patentin Ver. 2.0 - beta
; SEQ ID NO 17
; LENGTH: 5194
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:plasmid
; OTHER INFORMATION: IFF2B/XsupF4H
US-08-844-274-17

Query Match      10.4%; Score 73.4; DB 3; Length 5194;
Best Local Similarity 87.9%; Pred. No. 3.1e-13;
Matches 80; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 615 CTGCAGAACACGCGAGCTAGATTAAACCTAGAAAGATAATCATATTGTGACGTGTTAAA 674
Db 4773 CTCTAGAGGGATCCTCTAGATTAAACCTAGAAAGATAATCATATTGTGACGTGTTAAA 4714

QY 675 GATAATCATCGGTAAATAATTGACGCATGGGAT 705
Db 4713 GATAATCATCGGTAAATAATTGACGCATGTGTT 4683

RESULT 7
US-09-598-421-17/c
; Sequence 17, Application US/09598421
; Patent No. 6551825
; GENERAL INFORMATION:
; APPLICANT: Fraser Jr., Malcom J.
; APPLICANT: Shirk, Paul D.
; APPLICANT: Elick, Teri A.
; APPLICANT: Perera, Omaththage
; TITLE OF INVENTION: PiggyBac Transposon-Based Genetic Transformation System
; TITLE OF INVENTION: for Insects
; FILE REFERENCE: 0148.96
; CURRENT APPLICATION NUMBER: US/09/598,421
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: 60/016,234
; PRIOR FILING DATE: 1996-04-19

; OTHER INFORMATION: Description of Artificial Sequence:plasmid
; OTHER INFORMATION: IFF2B/Xpuc18.1
US-09-598-421-16

Query Match      10.4%; Score 73.4; DB 4; Length 5194;
Best Local Similarity 83.2%; Pred. No. 2.3e-13;
Matches 84; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 605 GCTCGACACGCTCGAGAACACGCGAGCTAGATTAAACCTAGAAAGATAATCATATTGTGAC 664
Db 2955 GCTCGGTACCCGGGGATCGGATCTCTAGATTAAACCTAGAAAGATAATCATATTGTGAC 2896

QY 665 GTAGCTTAAAGATAATCATCGTAAATAATTGACGCATGGGAT 705
Db 2895 GTAGCTTAAAGATAATCATCGTAAATAATTGACGCATGTGTT 2855

RESULT 6
US-08-844-274-17/c
; Sequence 17, Application US/08844274B
; Patent No. 6218185
; GENERAL INFORMATION:
; APPLICANT: Fraser Jr., Malcom J.
; APPLICANT: Shirk, Paul D.
; APPLICANT: Elick, Teri A.
; APPLICANT: Perera, Omaththage
; TITLE OF INVENTION: PiggyBac Transposon-Based Genetic Transformation System
; TITLE OF INVENTION: for Insects
; FILE REFERENCE: 0148.96
; CURRENT APPLICATION NUMBER: US/08/844,274B
; CURRENT FILING DATE: 1997-04-18
; EARLIER APPLICATION NUMBER: 60/016,234
; EARLIER FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patentin Ver. 2.0 - beta
; SEQ ID NO 17
; LENGTH: 5194
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:plasmid
; OTHER INFORMATION: IFF2B/XsupF4H
US-08-844-274-17

Query Match      10.4%; Score 73.4; DB 3; Length 5194;
Best Local Similarity 87.9%; Pred. No. 3.1e-13;
Matches 80; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 615 CTGCAGAACACGCGAGCTAGATTAAACCTAGAAAGATAATCATATTGTGACGTGTTAAA 674
Db 4773 CTCTAGAGGGATCCTCTAGATTAAACCTAGAAAGATAATCATATTGTGACGTGTTAAA 4714

QY 675 GATAATCATCGGTAAATAATTGACGCATGGGAT 705
Db 4713 GATAATCATCGGTAAATAATTGACGCATGTGTT 4683

RESULT 7
US-09-598-421-17/c
; Sequence 17, Application US/09598421
; Patent No. 6551825
; GENERAL INFORMATION:
; APPLICANT: Fraser Jr., Malcom J.
; APPLICANT: Shirk, Paul D.
; APPLICANT: Elick, Teri A.
; APPLICANT: Perera, Omaththage
; TITLE OF INVENTION: PiggyBac Transposon-Based Genetic Transformation System
; TITLE OF INVENTION: for Insects
; FILE REFERENCE: 0148.96
; CURRENT APPLICATION NUMBER: US/09/598,421
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: 60/016,234
; PRIOR FILING DATE: 1996-04-19

; OTHER INFORMATION: Description of Artificial Sequence:plasmid
; OTHER INFORMATION: IFF2B/XsupF4H
US-09-598-421-17

Query Match      10.4%; Score 73.4; DB 4; Length 5194;
Best Local Similarity 87.9%; Pred. No. 3.1e-13;
Matches 80; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 615 CTGCAGAACACGCGAGCTAGATTAAACCTAGAAAGATAATCATATTGTGACGTGTTAAA 674
Db 4773 CTCTAGAGGGATCCTCTAGATTAAACCTAGAAAGATAATCATATTGTGACGTGTTAAA 4714

QY 675 GATAATCATCGGTAAATAATTGACGCATGGGAT 705
Db 4713 GATAATCATCGGTAAATAATTGACGCATGTGTT 4683

RESULT 8
US-08-844-274-10
; Sequence 10, Application US/08844274B
; Patent No. 6218185
; GENERAL INFORMATION:
; APPLICANT: Fraser Jr., Malcom J.
; APPLICANT: Shirk, Paul D.
; APPLICANT: Elick, Teri A.
; APPLICANT: Perera, Omaththage
; TITLE OF INVENTION: PiggyBac Transposon-Based Genetic Transformation System
; TITLE OF INVENTION: for Insects
; FILE REFERENCE: 0148.96
; CURRENT APPLICATION NUMBER: US/08/844,274B
; CURRENT FILING DATE: 1997-04-18
; EARLIER APPLICATION NUMBER: 60/016,234
; EARLIER FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patentin Ver. 2.0 - beta
; SEQ ID NO 10
; LENGTH: 5679
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:plasmid
; OTHER INFORMATION: TRL
US-08-844-274-10

Query Match      9.6%; Score 67.8; DB 3; Length 5679;
Best Local Similarity 97.2%; Pred. No. 2.3e-11;
Matches 69; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 635 TTAAACCTAGAAAGATAATCATATTGTGACGTGTTAAAAGATAATCATCGGTGTTAAAATTG 694
Db 1080 TTAAACCTAGAAAGATAATCATATTGTGACGTGTTAAAAGATAATCATCGGTGTTAAAATTG 1139

QY 695 ACGCATGGGAT 705
Db 1140 ACGCATGTGTT 1150

RESULT 9
US-09-598-421-10
; Sequence 10, Application US/09598421
; Patent No. 6551825
; GENERAL INFORMATION:
; APPLICANT: Fraser Jr., Malcom J.
; APPLICANT: Shirk, Paul D.
; APPLICANT: Elick, Teri A.
; APPLICANT: Perera, Omaththage
```

```
; TITLE OF INVENTION: PiggyBac Transposon-Based Genetic Transformation System
; TITLE OF INVENTION: for Insects
; FILE REFERENCE: 0148.96
; CURRENT APPLICATION NUMBER: US/09/598,421
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: 60/016,234
; PRIOR FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO 10
; LENGTH: 5679
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:p3el.2 delta
; OTHER INFORMATION: TRL
US-09-598-421-10

Query Match          9.6%; Score 67.8; DB 4; Length 5679;
Best Local Similarity 97.2%; Pred. No. 2.5e-11;
Matches 69; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 635 TTAACCCCTAGAAAGATAATCATATTGTGACGTAGCTTTAAAGATAATCATCGGTAAATTTG 694
Db 1080 TTAACCCCTAGAAAGATAATCATATTGTGACGTAGCTTTAAAGATAATCATCGGTAAATTTG 1139

QY 695 ACGCATGGGAT 705
Db 1140 ACGCATGTGTT 1150

RESULT 10
US-08-844-274-13
; Sequence 13, Application US/08844274B
; Patent No. 6218185
; GENERAL INFORMATION:
; APPLICANT: Fraser Jr., Malcom J.
; APPLICANT: Shirk, Paul D.
; APPLICANT: Elick, Teri A.
; APPLICANT: Perera, Omathhage
; TITLE OF INVENTION: PiggyBac Transposon-Based Genetic Transformation System
; TITLE OF INVENTION: for Insects
; FILE REFERENCE: 0148.96
; CURRENT APPLICATION NUMBER: US/08/844,274B
; CURRENT FILING DATE: 1997-04-18
; EARLIER APPLICATION NUMBER: 60/016,234
; EARLIER FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO 13
; LENGTH: 6723
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Clone p3el.2H/S
US-08-844-274-13

Query Match          9.6%; Score 67.8; DB 3; Length 6723;
Best Local Similarity 97.2%; Pred. No. 2.5e-11;
Matches 69; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 635 TTAACCCCTAGAAAGATAATCATATTGTGACGTAGCTTTAAAGATAATCATCGGTAAATTTG 694
Db 1080 TTAACCCCTAGAAAGATAATCATATTGTGACGTAGCTTTAAAGATAATCATCGGTAAATTTG 1139

QY 695 ACGCATGGGAT 705
Db 1140 ACGCATGTGTT 1150

RESULT 11
US-08-844-274-14/c
; Sequence 14, Application US/08844274B
```

```
; Patent No. 6218185
; GENERAL INFORMATION:
; APPLICANT: Fraser Jr., Malcom J.
; APPLICANT: Shirk, Paul D.
; APPLICANT: Elick, Teri A.
; APPLICANT: Perera, Omathhage
; TITLE OF INVENTION: PiggyBac Transposon-Based Genetic Transformation System
; TITLE OF INVENTION: for Insects
; FILE REFERENCE: 0148.96
; CURRENT APPLICATION NUMBER: US/08/844,274B
; CURRENT FILING DATE: 1997-04-18
; EARLIER APPLICATION NUMBER: 60/016,234
; EARLIER FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO 14
; LENGTH: 6723
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:piggyBac/opd
US-08-844-274-14

Query Match          9.6%; Score 67.8; DB 3; Length 6723;
Best Local Similarity 97.2%; Pred. No. 2.5e-11;
Matches 69; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 635 TTAACCCCTAGAAAGATAATCATATTGTGACGTAGCTTTAAAGATAATCATCGGTAAATTTG 694
Db 5644 TTAACCCCTAGAAAGATAATCATATTGTGACGTAGCTTTAAAGATAATCATCGGTAAATTTG 5585

QY 695 ACGCATGGGAT 705
Db 5584 ACGCATGTGTT 5574

RESULT 12
US-09-598-421-13
; Sequence 13, Application US/09598421
; Patent No. 6551825
; GENERAL INFORMATION:
; APPLICANT: Fraser Jr., Malcom J.
; APPLICANT: Shirk, Paul D.
; APPLICANT: Elick, Teri A.
; APPLICANT: Perera, Omathhage
; TITLE OF INVENTION: PiggyBac Transposon-Based Genetic Transformation System
; TITLE OF INVENTION: for Insects
; FILE REFERENCE: 0148.96
; CURRENT APPLICATION NUMBER: US/09/598,421
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: 60/016,234
; PRIOR FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO 13
; LENGTH: 6723
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Clone p3el.2H/S
US-09-598-421-13

Query Match          9.6%; Score 67.8; DB 4; Length 6723;
Best Local Similarity 97.2%; Pred. No. 2.5e-11;
Matches 69; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 635 TTAACCCCTAGAAAGATAATCATATTGTGACGTAGCTTTAAAGATAATCATCGGTAAATTTG 694
Db 1080 TTAACCCCTAGAAAGATAATCATATTGTGACGTAGCTTTAAAGATAATCATCGGTAAATTTG 1139

QY 695 ACGCATGGGAT 705
Db 1140 ACGCATGTGTT 1150
```


RESULT 13
US-09-598-421-14/c
; Sequence 14, Application US/09598421
; Patent No. 6551825
; GENERAL INFORMATION:
; APPLICANT: Fraser Jr., Malcom J.
; APPLICANT: Shirk, Paul D.
; APPLICANT: Elick, Teri A.
; APPLICANT: Perera, Omaththage
; TITLE OF INVENTION: PiggyBac Transposon-Based Genetic Transformation System
; FILE REFERENCE: 0148.96
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: 60/016,234
; PRIOR FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO 14
; LENGTH: 6723
; TYPE: DNA
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: Description of Artificial Sequence: piggyBac/opd
US-09-598-421-14

Query Match 9.6%; Score 67.8; DB 4; Length 6723;
Best Local Similarity 97.2%; Pred. No. 2.5e-11;
Matches 69; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 635 TTAACCCCTAGAAAGATAATCATATTGTGACGTACGTTAAAGATAATCATCGTAAATTTG 694
Db 5644 TTAACCCCTAGAAAGATAATCATATTGTGACGTACGTTAAAGATAATCATCGTAAATTTG 5585

QY 695 ACGCATGGGAT 705
Db 5584 ACGCATGTGTT 5574

RESULT 14
US-08-844-274-20/c
; Sequence 20, Application US/08844274B
; Patent No. 6218185
; GENERAL INFORMATION:
; APPLICANT: Fraser Jr., Malcom J.
; APPLICANT: Shirk, Paul D.
; APPLICANT: Elick, Teri A.
; APPLICANT: Perera, Omaththage
; TITLE OF INVENTION: PiggyBac Transposon-Based Genetic Transformation System
; FILE REFERENCE: 0148.96
; CURRENT APPLICATION NUMBER: US/08/844,274B
; CURRENT FILING DATE: 1997-04-18
; EARLIER APPLICATION NUMBER: 60/016,234
; EARLIER FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO 20
; LENGTH: 7560
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: p3E1.2hs/opd
US-08-844-274-20

Query Match 9.6%; Score 67.8; DB 3; Length 7560;
Best Local Similarity 97.2%; Pred. No. 2.7e-11;
Matches 69; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 635 TTAACCCCTAGAAAGATAATCATATTGTGACGTACGTTAAAGATAATCATCGTAAATTTG 694

Db 4082 TTAACCCCTAGAAAGATAATCATATTGTGACGTACGTTAAAGATAATCATCGTAAATTTG 4023

QY 695 ACGCATGGGAT 705
Db 4022 ACGCATGTGTT 4012

RESULT 15
US-09-598-421-20/c
; Sequence 20, Application US/09598421
; Patent No. 6551825
; GENERAL INFORMATION:
; APPLICANT: Fraser Jr., Malcom J.
; APPLICANT: Shirk, Paul D.
; APPLICANT: Elick, Teri A.
; APPLICANT: Perera, Omaththage
; TITLE OF INVENTION: PiggyBac Transposon-Based Genetic Transformation System
; FILE REFERENCE: 0148.96
; CURRENT APPLICATION NUMBER: US/09/598,421
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: 60/016,234
; PRIOR FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO 20
; LENGTH: 7560
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: p3E1.2hs/opd
US-09-598-421-20

Query Match 9.6%; Score 67.8; DB 4; Length 7560;
Best Local Similarity 97.2%; Pred. No. 2.7e-11;
Matches 69; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 635 TTAACCCCTAGAAAGATAATCATATTGTGACGTACGTTAAAGATAATCATCGTAAATTTG 694
Db 4082 TTAACCCCTAGAAAGATAATCATATTGTGACGTACGTTAAAGATAATCATCGTAAATTTG 4023

QY 695 ACGCATGGGAT 705
Db 4022 ACGCATGTGTT 4012

Search completed: June 5, 2005, 23:07:32
Job time : 122.806 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 5, 2005, 13:38:04 ; Search time 2067.76 Seconds
(without alignments)
13014.775 Million cell updates/sec

Title: US-10-826-523-40

Perfect score: 707

Sequence: 1 ggaaccatgcgcgaatttt.....aaattgacgatggatcc 707

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gse1:*
9: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
C 1	585.4	82.8	682	7	CK781302	CK781302 UI-M-GIO-
C 2	585.4	82.8	703	7	CF537771	CF537771 UI-M-GIO-
C 3	585.4	82.8	711	7	CF743678	CF743678 UI-M-GIO-
C 4	585.4	82.8	853	9	CR065566	CR065566 Forward s
C 5	584.4	82.7	703	5	BQ154655	BQ154655 NF087C061
C 6	584.4	82.7	719	6	CD350897	CD350897 UI-M-GIO-
C 7	583.8	82.6	751	9	CR087413	CR087413 Forward s
C 8	580.4	82.1	716	9	CR131675	CR131675 Forward s
C 9	580	82.0	786	9	CR108810	CR108810 Forward s
C 10	565.4	80.0	691	6	CD350776	CD350776 UI-M-GIO-
C 11	564	79.8	762	9	CR035207	CR035207 Reverse s
C 12	563.4	79.7	749	6	CD351273	CD351273 UI-M-GIO-
C 13	562	79.5	733	9	CR139473	CR139473 Forward s
C 14	555	78.5	633	9	CR167678	CR167678 Reverse s
C 15	554.4	78.4	791	6	CB520716	CB520716 UI-M-GIO-
C 16	550.4	77.9	630	9	CR013093	CR013093 Reverse s
C 17	549	77.7	702	9	CR084987	CR084987 Reverse s
C 18	545.4	77.1	611	4	B1423083	B1423083 EST533749
C 19	543.4	76.9	716	7	CK781284	CK781284 UI-M-GIO-
C 20	540.8	76.5	716	7	CF851544	CF851544 pMA007xC
C 21	537.2	76.0	706	1	AV731514	AV731514 AV731514
C 22	535.4	75.7	730	9	CR077673	CR077673 Reverse s
C 23	525	74.3	635	9	BX982794	BX982794 Forward s
C 24	521.8	73.8	536	9	CR026633	CR026633 Reverse s

C 25	516	73.0	887	9	CR192748	CR192748 Forward s
C 26	515	72.8	885	9	CR255010	CR255010 Forward s
C 27	498.4	70.5	683	6	CD348896	CD348896 UI-M-FY0-
C 28	493	69.7	495	4	BG457760	BG457760 NF036R03P
C 29	490.4	69.4	786	6	CB520098	CB520098 UI-M-GIO-
C 30	482.6	68.3	795	6	CD351738	CD351738 UI-M-GIO-
C 31	477.4	67.5	760	7	CF538326	CF538326 UI-M-GIO-
C 32	474.4	67.1	620	9	CR034585	CR034585 Forward s
C 33	474.4	67.1	620	9	CR086282	CR086282 Reverse s
C 34	474.4	67.1	620	9	CR171547	CR171547 Reverse s
C 35	471.8	66.7	620	9	CR078862	CR078862 Forward s
C 36	467.4	66.1	752	9	CR088896	CR088896 Forward s
C 37	459.6	65.0	805	6	CD352530	CD352530 UI-M-GIO-
C 38	457	64.6	676	5	BQ154673	BQ154673 NF095C111
C 39	455.4	64.4	483	9	CR068828	CR068828 Reverse s
C 40	447.4	63.3	547	9	CR148906	CR148906 Forward s
C 41	447.4	63.3	885	9	CR024214	CR024214 Reverse s
C 42	444.4	62.9	599	9	CR079016	CR079016 Forward s
C 43	443.4	62.7	702	9	CR038019	CR038019 Forward s
C 44	440.8	62.3	538	9	CR029330	CR029330 Forward s
C 45	437.6	61.9	901	6	CA328434	CA328434 UI-M-FY0-

ALIGNMENTS

RESULT 1
CK781302/c
LOCUS CK781302 682 bp mRNA linear EST 23-FEB-2004
DEFINITION UI-M-GIO-clp-g-05-0-UI-ri NIH BMAP_GIO Mus musculus CDNA clone
IMAGE:30613060 5', mRNA sequence.
ACCESSION CK781302
VERSION CK781302.1 GI:42746980
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 682)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/mousefl.html
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

FEATURES

Seq primer: PYX-5.
Location/Qualifiers
1. .682
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30613060"
/tissue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH BMAP_GIO"
/note="Organ: Brain; Vector: PYX- Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
Bonatolo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into PYX-Asc vector. The library tag

sequence located between the Not I site and the polyA tail is AGGACACAG. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN		FEATURES		Location/Qualifiers		source	
Query Match		82.8%; Score 585.4; DB 7; Length 682;		1. .703		/organism="Mus musculus"	
Best Local Similarity		99.8%; Pred. No. 2.4e-164;		/mol_type="mRNA"		/strain="CS7BL/6"	
Matches 586; Conservative 0; Mismatches 1; Indels 0; Gaps 0;				/db_xref="taxon:10090"		/clone="IMAGE:30537799"	
QY 45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCGGCTCAGTTCATCGCCCAAGC 104				/tissue_type="whole brain"		/dev_stages="embryo 13.5,14.5,16.5,17.5dpc"	
Db 611 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCGGCTCAGTTCATCGCCCAAGC 552				/lab_host="DH10B (T1 phage resistant)"		/clone_lib="NIH_BMAP_G10"	
QY 105 TGGCGCTATCTGGGCATCGGGAGGAAGCCCGTGCCTTTTCCCGCAGGTTGAAGCG 164				/note="Organ: Brain; Vector: pYX- Asc; Site:1: EcoR I; Site 2: Not I; The library was constructed according			
Db 551 TGGCGCTATCTGGGCATCGGGAGGAAGCCCGTGCCTTTTCCCGCAGGTTGAAGCG 492				Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose			
QY 165 GCATGAAAGAGTTTCCCGAGGATGACTGCTGCTGATTGAGTTTCAGCGGAAACGCACG 224				gel. First strand cDNA synthesis was primed with oligo-dT			
Db 491 GCATGAAAGAGTTTCCCGAGGATGACTGCTGCTGATTGAGTTTCAGCGGAAACGCACG 432				primer containing a Not I site. Double strand cDNA was			
QY 225 TTTTACCATGATGATTCGGGAAGGTGTGGCCATGACGCGCTTTTAAACGGTGAACCTGTTTCGTT 284				size selected according to mRNA size fraction, ligated			
Db 431 TTTTACCATGATGATTCGGGAAGGTGTGGCCATGACGCGCTTTTAAACGGTGAACCTGTTTCGTT 372				with EcoR I adaptor, digested with NotI and then cloned			
QY 285 CAGGCCACCTGGGATACCAAGTTCGCGGCTTTTCCGACACAGTTCGGATGTCGTCAGC 344				directionally into pYX-Asc vector. The library tag			
Db 371 CAGGCCACCTGGGATACCAAGTTCGCGGCTTTTCCGACACAGTTCGGATGTCGTCAGC 312				sequence located between the Not I site and the polyA tail			
QY 345 CCGAAGCGCATCAGCAACCCGCAACATACCGCGACAGCGGAACCTGCGTGCCTGTCG 404				is AGGACACAG. This library was created for the University			
Db 311 CCGAAGCGCATCAGCAACCCGCAACATACCGCGACAGCGGAACCTGCGTGCCTGTCG 252				of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the			
QY 405 CAGATTAATGACAGCGGTGCGGCTGGGATATTACGTGAGGAGACGGGTATCTCTGCG 464				Developing Mouse Nervous System', supported by National			
Db 251 CAGATTAATGACAGCGGTGCGGCTGGGATATTACGTGAGGAGACGGGTATCTCTGCG 192				Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,			
QY 465 TGGATCCCGAGAAATGGACATGATACCCGTGAGTTACCCGGCGGCGCGCTCGTTC 524				program coordinator."			
Db 191 TGGATCCCGAGAAATGGACATGATACCCGTGAGTTACCCGGCGGCGCGCTCGTTC 132							
QY 525 ATTACAGTTTTTGAACCCGTGGAGGAGCGGACACTCGCGTGCCTGCAATGTTTTACAGC 584							
Db 131 ATTACAGTTTTTGAACCCGTGGAGGAGCGGACACTCGCGTGCCTGCAATGTTTTACAGC 72							
QY 585 GTGATGGAGCAGATGAAGATGCTCGACAGCTGCGAACAACGCGAGCT 631							
Db 71 GTGATGGAGCAGATGAAGATGCTCGACAGCTGCGAACAACGCGAGCT 25							

RESULT 2
CF537771/c
LOCUS
DEFINITION
IMAGE:30537799 5', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

CF537771 703 bp mRNA linear EST 12-SEP-2003
UI-M-GIO-chn-m-08-0-UI-r1 NIH_BMAP_G10 Mus musculus cDNA clone
IMAGE:30537799 5', mRNA sequence.
CF537771
EST.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 703)
NIH-MGC
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA library preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mousefl.html>
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

Seq primer: pYX-5.
Location/Qualifiers
1. .703
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CS7BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30537799"
/tissue_type="whole brain"
/dev_stages="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_G10"
/note="Organ: Brain; Vector: pYX- Asc; Site:1: EcoR I; Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is AGGACACAG. This library was created for the University
of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator."

ORIGIN

Query Match	82.8%;	Score	585.4;	DB	7;	Length	703;
Best Local Similarity	99.8%;	Pred. No.	2.4e-164;				
Matches	586;	Conservative	0;	Mismatches	1;	Indels	0;
Gaps	0;						
QY	45	ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCGGCTCAGTTCATCGCCCAAGC	104				
Db	643	ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCGGCTCAGTTCATCGCCCAAGC	584				
QY	105	TGGCGCTATCTGGGCATCGGGAGGAAGAAGCCCGTGCCTTTTCCCGCAGGTTGAAGCG	164				
Db	583	TGGCGCTATCTGGGCATCGGGAGGAAGAAGCCCGTGCCTTTTCCCGCAGGTTGAAGCG	524				
QY	165	GCATGGAAGAGTTTCCCGAGGATGACTGCTGCTGCATTGACGTTGACGGAACCGCACG	224				
Db	523	GCATGGAAGAGTTTCCCGAGGATGACTGCTGCTGCATTGACGTTGACGGAACCGCACG	464				
QY	225	TTTTACCATGATGATTCGGGAAGGTGTGGCCATGACGCGCTTTTAAACGGTGAACCTGTCGTT	284				
Db	463	TTTTACCATGATGATTCGGGAAGGTGTGGCCATGACGCGCTTTTAAACGGTGAACCTGTCGTT	404				
QY	285	CAGGCCACCTGGGATACCAAGTTGTCGGGCTTTTCCGGAACACAGTTCGGGATGTCGTCAGC	344				
Db	403	CAGGCCACCTGGGATACCAAGTTGTCGGGCTTTTCCGGAACACAGTTCGGGATGTCGTCAGC	344				
QY	345	CCGAAGCGCATCAGCAACCCGCAACATAACCGCGACAGCCGGAACTGCGGTGCCGTGTG	404				
Db	343	CCGAGCGCATCAGCAACCCGCAACATAACCGCGACAGCCGGAACTGCGGTGCCGTGTG	284				
QY	405	CAGATTAATGACAGCGGTGCGGCTGGGATATTACGTGACGAGGAACGGGTATCTCTGCG	464				
Db	283	CAGATTAATGACAGCGGTGCGGCTGGGATATTACGTGACGAGGACGGGTATCTCTGCG	224				
QY	465	TGGATCCCGCAGAAATGGACATGATACCCCGTGAAGTTTACCCCGCGCGCGCTCGTTC	524				
Db	223	TGGATCCCGCAGAAATGGACATGATACCCCGTGAAGTTTACCCCGCGCGCGCTCGTTC	164				
QY	525	ATTACAGTTTTTGAACCCGTGGAGGACGGCAGACTCGCGTGCCTGCAATGTTTTTACAGC	584				

QY 165 GCATGGAAGAGTTTCCGAGGATGACTGCTGCTGATTGACGTTGAGCGAAACGACG 224
DB 384 GCATGGAAGAGTTTCCGAGGATGACTGCTGCTGATTGAGTTGAGCGAAACGACG 443
QY 225 TTATCATGATGATTGGGAAGGTGGCCATGACGCTTTAAACGGTGAACCTGTCGTT 284
DB 444 TTATCATGATGATTGGGAAGGTGGCCATGACGCTTTAAACGGTGAACCTGTCGTT 503
QY 285 CAGGCCACCTGGATACCAAGTTCTGCGCGCTTTCCGACACAGTTCCGGATGTCACG 344
DB 504 CAGGCCACCTGGATACCAAGTTCTGCGCGCTTTCCGACACAGTTCCGGATGTCACG 563
QY 345 CCGAAGCGCATCAGCAACCCGCAACATACCGCGACAGCGCAACTGCCGTCGCCGTGTG 404
DB 564 CCGAAGCGCATCAGCAACCCGCAACATACCGCGACAGCGCAACTGCCGTCGCCGTGTG 623
QY 405 CAGATTAAATGACAGCGGTGGCGCTGGGATATTAGTTCAGGAGACGGGTATCTGCG 464
DB 624 CAGATTAAATGACAGCGGTGGCGCTGGGATATTAGTTCAGGAGACGGGTATCTGCG 683
QY 465 TGGATGCCGACAGAAATGGACATGGATACCCCTGAGTTACCCGCGCGCGCTGCTTC 524
DB 684 TGGATGCCGACAGAAATGGACATGGATACCCCTGAGTTACCCGCGCGCGCTGCTTC 743
QY 525 ATTACAGTTTGTGAACCCCTGGAGACGGCGAGACTCGCGGTGCAAAATGTGTTTACAGC 584
DB 744 ATTACAGTTTGTGAACCCCTGGAGACGGCGAGACTCGCGGTGCAAAATGTGTTTACAGC 803
QY 585 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGACGCT 631
DB 804 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGACGCT 850

RESULT 5
BQ154655/c
LOCUS
DEFINITION
NF087C06IR1F1050 Irradiated Medicago truncatula cdna clone
EST 24-APR-2002
BQ154655
BQ154655.1 GI:20291714
EST.
Medicago truncatula (barrel medic)

ORGANISM
Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliales;
Medicago.

REFERENCE
1 (bases 1 to 703)
Torres-Jerez,I., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J.,
Flores,H.R., Inman,J.T., Weller,J.W. and May,G.D.
Expressed Sequence Tags from the Samuel Roberts Noble Foundation
Medicago truncatula irradiated library
Unpublished (2001)
Contact: May GD

JOURNAL
COMMENT
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 224 6650
Fax: 580 224 6692
Email: gdmay@noble.org
Insert Length: 703 Std Error: 0.00
Plate: 087 row: C column: 06
Seq primer: TCACACGAAACAGCTATGAC.
Location/Qualifiers
1. 703

FEATURES
source
/organism="Medicago truncatula"
/mol_type="mRNA"
/db_xref="taxon:3880"
/clone="NF087C06IR"
/tissue_type="seedlings"
/dev_stage="seedling"
/clone_lib="Irradiated"

/note="Vector: Lambda Zap: Seedlings were exposed either to 100 Gy gamma or 0.5, 1, 5, or 10 kJ/m2 UV irradiation. Gamma-irradiated samples were harvested at 6, 12, 24 and 48 hours after treatment. UV-irradiated samples were harvested 24 hours post-treatment. cDNA was prepared from polyA+ enriched, pooled samples of equivalent amounts of total RNA from each sample. The cDNA was directionally ligated into the Uni-Zap XR vector (Stratagene) and packaged using the Gigapack III Gold packaging extracts. Phagemids containing cDNA inserts were in vivo excised from the recombinant Uni-Zap XR vector using EXaseIst helper phage and the E. coli strain XL1-Blue MRP' (Stratagene). Excised plasmids were plated using SOLR cells."

ORIGIN

Query Match 82.7%; Score 584.4; DB 5; Length 703;
Best Local Similarity 99.7%; Pred. No. 4.8e-164;
Matches 585; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTCCGGCTCAGTCATCGCCCAAGC 104
DB 624 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTCCGGCTCAGTCATCGCCCAAGC 565
QY 105 TGGCGCTATCTGGGCATCGGGAGGAAGAACCCGTCGCTTTTCCGGGAGGTTGAAGCG 164
DB 564 TGGCGCTATNTGGGCATCGGGAGGAAGAACCCGTCGCTTTTCCGGGAGGTTGAAGCG 505
QY 165 GCATGGAAGAGTTTCCGAGGATGACTGCTGCTGCATTGAGTTGAGGGAACGACG 224
DB 504 GCATGGAAGAGTTTCCGAGGATGACTGCTGCTGCATTGAGTTGAGGGAACGACG 445
QY 225 TTTTACCATGATGATTCCGGGAAGGTGTGGCATGCAAGCTTTAAACGGTGAACCTGTCGTT 284
DB 444 TTTTACCATGATGATTCCGGGAAGGTGTGGCATGCAAGCTTTAAACGGTGAACCTGTCGTT 385
QY 285 CAGGCCACCTGGGATACCAAGTTTCGTCGGGCTTTTCCGACACAGTTCCGGATGTCACG 344
DB 384 CAGGCCACCTGGGATACCAAGTTTCGTCGGGCTTTTCCGACACAGTTCCGGATGTCACG 325
QY 345 CCGAAGCGCATCAGCAACCCGCAACATACCGCGACAGCGCAACTGCCGTCGCCGTGTG 404
DB 324 CCGAAGCGCATCAGCAACCCGCAACATACCGCGACAGCGCAACTGCCGTCGCCGTGTG 265
QY 405 CAGATTAAATGACAGCGGTGGCGCTGGGATATTAGTTCAGGAGACGGGTATCTGCGC 464
DB 264 CAGATTAAATGACAGCGGTGGCGCTGGGATATTAGTTCAGGAGACGGGTATCTGCGC 205
QY 465 TGGATGCCGCAAAATGGACATGGATACCCGTCGAGTTACCCGCGCGCGCTGCTGTTTC 524
DB 204 TGGATGCCGCAAAATGGACATGGATACCCGTCGAGTTACCCGCGCGCGCTGCTGTTTC 145
QY 525 ATTACAGTTTGTGAACCCCTGGAGACGGCGAGACTCGCGGTGCAAAATGTGTTTACAGC 584
DB 144 ATTACAGTTTGTGAACCCCTGGAGACGGCGAGACTCGCGGTGCAAAATGTGTTTACAGC 85
QY 585 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGACGCT 631
DB 84 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGACGCT 38

RESULT 6

CD350897/c
LOCUS
DEFINITION
UI-M-G10-cgh-a-18-0-UI.r1 NIH_BMAP_G10 Mus musculus cdna clone
IMAGE:6853243 5', mRNA sequence.
CD350897
CD350897.1 GI:31142412
EST.
Mus musculus (house mouse)
Mus musculus
ORGANISM

CD350897 719 bp mRNA linear EST 09-JUL-2003
UI-M-G10-cgh-a-18-0-UI.r1 NIH_BMAP_G10 Mus musculus cdna clone
IMAGE:6853243 5', mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus
Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 719)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-r@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
cDNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mousefl.html>
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

Seq primer: pYX-5.

FEATURES

Source

Location/Qualifiers
1..719
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6853243"
/tissue_type="whole brain"
/dev_stage="embryo 13.5 14.5 16.5 17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH BMAP_G10"
/note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is AGCGAGACAG. This library was created for the University
Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator."

ORIGIN

Query Match 82.7%; Score 584.4; DB 6; Length 719;
Best Local Similarity 99.7%; Pred. No. 4.8e-164;
Matches 585; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGGTCTAGTCATCGCCCAAGC 104
DB 663 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGGTCTAGTCATCGCCCAAGC 604
QY 105 TGGCGCTATCTGGGCATCGGGAGGAAGACCGCGTCTTTTTCGGGTCTAGTCATCGCCCAAGC 164
DB 603 TGGCGCTATCTGGGCATCGGGAGGAAGACCGCGTCTTTTTCGGGTCTAGTCATCGCCCAAGC 544
QY 165 GCATGAAAGAGTTTCCCGAGGATGACTGCTGCTGATTCGTCATTCGAGCGAAGAAACGACG 224
DB 543 GCATGAAAGAGTTTCCCGAGGATGACTGCTGCTGATTCGTCATTCGAGCGAAGAAACGACG 484
QY 225 TTATACCATGATGATTCGGGAAGGTGGCCATGTCAGCCCTTTTAAACGGTGAACCTGTCGTT 284
DB 483 TTATACCATGATGATTCGGGAAGGTGGCCATGTCAGCCCTTTTAAACGGTGAACCTGTCGTT 424
QY 285 CAGGCCACCTGGGATACCAAGTTCGTCGGGTCTTTTCGGGACACAGTTCCGGATGGTCAGC 344
DB 423 CAGGCCACCTGGGATACCAAGTTCGTCGGGTCTTTTCGGGACACAGTTCCGGATGGTCAGC 364
QY 345 CCGAAGCGCATCAGCAACCCGAAACAATACCGGCGACAGCCGGAACCTGCCGTGCCGTGTG 404
DB 363 CCGAAGCGCATCAGCAACCCGAAACAATACCGGCGACAGCCGGAACCTGCCGTGCCGTGTG 304
QY 405 CAGATTAAATGACACCGGTGCCGCTGGGATATTAGCTGACGAGGACCGGTATCTCTGC 464

DB 303 CAGATTAAATGACACCGGTGCCGCTGGGATATTAGCTGACGAGGACCGGTATCTCTGC 244
QY 465 TGGATGCCCGAGAAATGGACATGATACCCCGTAGTTACCCGCGGGCGCTCGTTC 524
DB 243 TGGATGCCCGAGAAATGGACATGATACCCCGTAGTTACCCGCGGGCGCTCGTTC 184
QY 525 ATTCAGGTTTGAACCCCGTAGGACGGGACGACTCGGGTGCRAATGTTTACACG 584
DB 183 ATTCAGGTTTGAACCCCGTAGGACGGGACGACTCGGGTGCRAATGTTTACACG 124
QY 585 GTGATGGACGACATGAAGATGCTCGACACGCTGCAGAAACGCGAGCT 631
DB 123 GTGATGGACGACATGAAGATGCTCGACACGCTGCAGAAACGCGAGCT 77

RESULT 7

CR087413/c

LOCUS CR087413 751 bp DNA linear GSS 05-JUL-2004
DEFINITION Forward strand read from insert in 3'HPRT insertion targeting and
chromosome engineering clone MHP438f24, genomic survey sequence.

ACCESSION CR087413

VERSION CR087413.1 GI:49821005

KEYWORDS GSS; Genome survey sequence; MICE.

SOURCE Mus musculus (house mouse)

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutelestomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 751)

AUTHORS Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,

Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,

Rogers,J. and Bradley,A.

Direct Submission

JOURNAL Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,

CB10 1SA, UK. <http://www.sanger.ac.uk/MICER>

FEATURES Location/Qualifiers

Source

1..751

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

/clone="MHP438f24"

/clone_lib="MHPp"

ORIGIN

Query Match 82.6%; Score 583.8; DB 9; Length 751;
Best Local Similarity 99.7%; Pred. No. 7.3e-164;
Matches 585; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGGTCTAGTCATCGCCCAAGC 104
DB 623 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGGTCTAGTCATCGCCCAAGC 564
QY 105 TGGCGCTATCTGGGCATCGGGAGGAAGACCGCGTCTTTTTCGGGTCTAGTCATCGCCCAAGC 164
DB 563 TGGCGCTATCTGGGCATCGGGAGGAAGACCGCGTCTTTTTCGGGTCTAGTCATCGCCCAAGC 504
QY 165 GCATGAAAGAGTTTCCCGAGGATGACTGCTGCTGATTCGTCATTCGAGCGAAGAAACGACG 224
DB 503 GCATGAAAGAGTTTCCCGAGGATGACTGCTGCTGATTCGTCATTCGAGCGAAGAAACGACG 444
QY 225 TTATACCATGATGATTCGGGAAGGTGGCCATGTCAGCCCTTTTAAACGGTGAACCTGTCGTT 284
DB 443 TTATACCATGATGATTCGGGAAGGTGGCCATGTCAGCCCTTTTAAACGGTGAACCTGTCGTT 384
QY 285 CAGGCCACCTGGGATACCAAGTTCGTCGGGTCTTTTCGGGACACAGTTCCGGATGGTCAGC 344
DB 383 CAGGCCACCTGGGATACCAAGTTCGTCGGGTCTTTTCGGGACACAGTTCCGGATGGTCAGC 324
QY 345 CCGAAGCGCATCAGCAACCCGAAACAATACCGGCGACAGCCGGAACCTGCCGTGCCGTGTG 404
DB 323 CCGAAGCGCATCAGCAACCCGAAACAATACCGGCGACAGCCGGAACCTGCCGTGCCGTGTG 264
QY 405 CAGATTAAATGACACCGGTGCCGCTGGGATATTAGCTGACGAGGACCGGTATCTCTGC 464

```
Db 263 CAGATTAAATGACAGCGGTGCGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 204
QY 465 TGGATGCCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGGGGGCGCCCTCGTTTC 524
Db 203 TGGATGCCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGGGGGCGCCCTCGTTTC 144
QY 525 ATTACAGTTTTTTGAACCCCGTGGAGGACGGCGAGACTCGCGGTGCAAAATGTGTTTTACAGC 584
Db 143 ATTACAGTTTTTTGAACCCCGTGGAGGACGGCGAGACTCGCGGTGCAAAATGTGTTTTACAGC 84
QY 585 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGAGCT 631
Db 83 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGAGCT 37

RESULT 8
CR131675/c
LOCUS
DEFINITION
Forward strand read from insert in 3'HPRT insertion targeting and
chromosome engineering clone MHP438f23, genomic survey sequence.
ACCESSION
CR131675.1 GI:49879128
VERSION
GSS; genome survey sequence; MICE.
KEYWORDS
Mus musculus (house mouse)
SOURCE
Mus musculus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 716)
AUTHORS
Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
Rogers,J. and Bradley,A.
Direct Submission
TITLE
Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
JOURNAL
CB10 13A, UK. http://www.sanger.ac.uk/MICER
FEATURES
Location/Qualifiers
source
1..716
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="MHP438f23"
/clone_lib="MHPP"

ORIGIN
Query Match 82.1%; Score 580.4; DB 9; Length 716;
Best Local Similarity 99.8%; Pred. No. 7.6e-163;
Matches 591; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGGTCTAGTCATCGCCCAAGC 104
Db 582 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGGTCTAGTCATCGCCCAAGC 523
QY 105 TGGCGCTATCTGGGCATCGGGAGGAGAGCCCGTCCCTTTTCCCGGAGGTTGAAGCG 164
Db 522 TGGCGCTATCTGGGCATCGGGAGGAGAGCCCGTCCCTTTTCCCGGAGGTTGAAGCG 463
QY 165 GCATGGAAGAGTTTCCCGAGGATGACTGCTGTCGATTGAGCTTTGAGCGAAAACGACG 224
Db 462 GCATGGAAGAGTTTCCCGAGGATGACTGCTGTCGATTGAGCTTTGAGCGAAAACGACG 403
QY 225 TTATCCATGATGATTGGGAAAGGTGTGGCCATGACGACCTTTTAAACGGTGAACTGTTCTGTT 284
Db 402 TTATCCATGATGATTGGGAAAGGTGTGGCCATGACGACCTTTTAAACGGTGAACTGTTCTGTT 343
QY 285 CAGGCCACTCGGATACAGTTGCTCGCGCTTTTCCGACACAGTTCCGGATGCTCAGC 344
Db 342 CAGGCCACTCGGATACAGTTGCTCGCGCTTTTCCGACACAGTTCCGGATGCTCAGC 283
QY 345 CCGAAGCGCATCAGCAACCCGAAACAAATACCGCGCAGCAGCGGAACCTGCGGTGCCGTGTG 404
Db 282 CCGAAGCGCATCAGCAACCCGAAACAAATACCGCGCAGCAGCGGAACCTGCGGTGCCGTGTG 223
QY 405 CAGATTAAATGACAGCGGTGCGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 464
```

```
Db 222 CAGATTAAATGACAGCGGTGCGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 163
QY 465 TGGATGCCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGGGGGCGCCCTCGTTTC 524
Db 162 TGGATGCCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGGGGGCGCCCTCGTTTC 103
QY 525 ATTACAGTTTTTTGAACCCCGTGGAGGACGGCGAGACTCGCGGTGCAAAATGTGTTTTACAGC 584
Db 102 ATTACAGTTTTTTGAACCCCGTGGAGGACGGCGAGACTCGCGGTGCAAAATGTGTTTTACAGC 43
QY 585 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACG 626
Db 42 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACG 1

RESULT 9
CR108810/c
LOCUS
DEFINITION
Forward strand read from insert in 3'HPRT insertion targeting and
chromosome engineering clone MHP22016, genomic survey sequence.
ACCESSION
CR108810
VERSION
CR108810.1 GI:49856225
KEYWORDS
GSS; genome survey sequence; MICE.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 786)
AUTHORS
Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
Rogers,J. and Bradley,A.
Direct Submission
TITLE
Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
JOURNAL
CB10 13A, UK. http://www.sanger.ac.uk/MICER
FEATURES
Location/Qualifiers
source
1..786
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="MHP22016"
/clone_lib="MHPP"

ORIGIN
Query Match 82.0%; Score 580; DB 9; Length 786;
Best Local Similarity 100.0%; Pred. No. 1e-162;
Matches 580; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 TCATCAGGATCATATCGTCGGGTCTTTTTCGGGTCTAGTCATCGCCCAAGCTGCGCT 111
Db 786 TCATCAGGATCATATCGTCGGGTCTTTTTCGGGTCTAGTCATCGCCCAAGCTGCGCT 727
QY 112 ATCTGGGCATCGGGAGGAGAAAGCCCGTCCCTTTTCCCGGAGTTGAAGCGGATGGA 171
Db 726 ATCTGGGCATCGGGAGGAGAAAGCCCGTCCCTTTTCCCGGAGTTGAAGCGGATGGA 667
QY 172 AAGAGTTTCCCGAGGATGACTGCTGTCGATTGAGCTTTGAGCGAAAACGACGTTTACCA 231
Db 666 AAGAGTTTCCCGAGGATGACTGCTGTCGATTGAGCTTTGAGCGAAAACGACGTTTACCA 607
QY 232 TGATGATTCCGGAAGGTTGGCCATGCACGCTTTTAAACGGTGAACTGTTCTGTTACGCCCA 291
Db 606 TGATGATTCCGGAAGGTTGGCCATGCACGCTTTTAAACGGTGAACTGTTCTGTTACGCCCA 547
QY 292 CCTGGGATACCAAGTTCTGTCGGGCTTTTCCGACACAGTTCCGGATGTTCCAGCCGAAGC 351
Db 546 CTTGGGATACCAAGTTCTGTCGGGCTTTTCCGACACAGTTCCGGATGTTCCAGCCGAAGC 487
QY 352 GCATCAGCAACCCGAAACAAATACCGCGCAGCAGCGGAACCTGCGGTGCCGTGTGCAGATTA 411
Db 486 GCATCAGCAACCCGAAACAAATACCGCGCAGCAGCGGAACCTGCGGTGCCGTGTGCAGATTA 427
QY 412 ATGACAGCGGTGCGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGCTGGATGC 471
```


Db 426 ATGACCGGGTCCGGCGCTGGGATATTACGTACAGGAGCAGGATTCCTCGCTGGTAGTC 367

Qy 472 CGCAGAAATGGACATGGATACCCCCTGAGTTACCCGGCGGGCGCTCGTTCAATTCACG 531

Db 366 CGCAGAAATGGACATGGATACCCCCTGAGTTACCCGGCGGGCGCTCGTTCAATTCACG 307

Qy 532 TTTTGAACCCGTGGAGCAGCGGCAGACTCCGGTGCAAAATGCTTTTTACAGCGTGATGG 591

Db 306 TTTTGAACCCGTGGAGCAGCGGCAGACTCCGGTGCAAAATGCTTTTTACAGCGTGATGG 247

Qy 592 AGCAGATGAAGATGCTCTGCACACGCTGCAGAACACGCCAGCT 631

Db 246 AGCAGATGAAGATGCTCTGCACACGCTGCAGAACACGCCAGCT 207

RESULT 10
CD350776/c

LOCUS CD350776 691 bp mRNA linear EST 09-JUL-2003

DEFINITION UI-M-GIO-cgh-g-19-0-UI.r1 NIH BMAP_GIO Mus musculus cDNA clone IMAGE:6853388 5', mRNA sequence.

ACCESSION CD350776

VERSION CD350776.1 GI:31142363

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 691)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)

JOURNAL Contact: Robert Strausberg, Ph.D.
Email: cqabbs@remail.nih.gov

COMMENT Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at <http://genome.uiowa.edu/distribution/mousef1.html>
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)
Seq primer: pYX-5.
Location/Qualifiers
1..691
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6853388"
/tissue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_GIO"
/notes="Organ: Brain; Vector: pYX-Asc; Site_1: Ecor I; Site_2: Not I; The library was constructed according to Ronald, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with Ecor I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is AGCAGACAG. This library was created for the University of Iowa Brain Anatomy Project (BMAP). 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Query Match 80.0%; Score 565.4; DB 6; Length 691;
Best Local Similarity 99.8%; Pred. No. 2.4e-158;
Matches 566; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	45	ATCTAGCTGCATCAGGATCATATCGTCGGGTCCTTTTTCGGGTCAGTCAATCGCCCAAGC	104
Db	567	ATCCAGCTGCATCAGGATCATATCGTCGGGTCCTTTTTCGGGTCAGTCAATCGCCCAAGC	508
Qy	105	TGCGCGTATCTCGGCATCGGGAGGAGAAAGCCGTCGCTTTTTCGGCGAGGTTGAAGCG	164
Db	507	TGCGCGTATCTCGGCATCGGGAGGAGAAAGCCGTCGCTTTTTCGGCGAGGTTGAAGCG	448
Qy	165	GCATGGAAGAGTTTTCGCCAGGATGACTGCTGCTGCATTGACGTTGAGCGAAACGCACG	224
Db	447	GCATGGAAGAGTTTTCGCCAGGATGACTGCTGCTGCATTGACGTTGAGCGAAACGCACG	388
Qy	225	TTTACCATGATGATTCGGGAAGGTGTGGCGCATGCACCGCTTTTAAACGCGTAACTGTTGCTT	284
Db	387	TTTACCATGATGATTCGGGAAGGTGTGGCGCATGCACCGCTTTTAAACGCGTAACTGTTGCTT	328
Qy	285	CAGGCGCACCTGGGATACCAAGTTCGTCGCGGCTTTTTCGGGACAGTTCGCGATGCTCAGC	344
Db	327	CAGGCGCACCTGGGATACCAAGTTCGTCGCGGCTTTTTCGGGACAGTTCGCGATGCTCAGC	268
Qy	345	CCGAAGCGCATCAGCAACCCGAACAATACCGGCGGACACGCGGAACTGCCGTCGCGGTGTG	404
Db	267	CCGAAGCGCATCAGCAACCCGAACAATACCGGCGGACACGCGGAACTGCCGTCGCGGTGTG	208
Qy	405	CAGATTAATGACAGCGGTTCGCGGCTGGGATATTACGTCAAGCAGGACGCGGTATCCTGGC	464
Db	207	CAGATTAATGACAGCGGTTCGCGGCTGGGATATTACGTCAAGCAGGACGCGGTATCCTGGC	148
Qy	465	TGATCGCGCAGAAATGACATGATACCCCGTAGTACCCCGCGGCGCGGCTCGTTC	524
Db	147	TGATCGCGCAGAAATGACATGATACCCCGTAGTACCCCGCGGCGCGGCTCGTTC	88
Qy	525	ATTACAGTTTTTCAACCCGTTGAGGACGGGACAGTTCGCGGTGCAAAATGTTTTACAGC	584
Db	87	ATTACAGTTTTTCAACCCGTTGAGGACGGGACAGTTCGCGGTGCAAAATGTTTTACAGC	28
Qy	585	GTGATGAGCAGATGAAGATGCTCGAC	611
Db	27	GTGATGAGCAGATGAAGATGCTCGAC	1
RESULT 11			
CR035207			
LOCUS			
DEFINITION			
CR035207 762 bp DNA linear GSS 05-JUL-2004			
Reverse strand read from insert in 5'HPRT insertion targeting and			
chromosome engineering clone MHPN278n13, genomic survey sequence.			
CR035207.1 GI:49768262			
VERSION			
KEYWORDS			
SOURCE			
MUS musculus (house mouse)			
ORGANISM			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
REFERENCE			
AUTHORS			
Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,			
Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,			
Rogers,J. and Bradley,A.			
Direct Submission			
TITLE			
JOURNAL			
Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,			
CB10 1SA, UK. http://www.sanger.ac.uk/MICER			
FEATURES			
source			
1..762			
Location/Qualifiers			
/organism="Mus musculus"			
/mol_type="genomic DNA"			
/db_xref="taxon:10090"			
/clone="MHPN278n13"			
/clone_lib="MHPN"			
ORIGIN			
Query Match 79.8%; Score 564; DB 9; Length 762;			
Best Local Similarity 100.0%; Pred. No. 6.5e-158;			
Matches 564; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			

```
QY 68 CGTCCGGCTCTTTTCCCGGCTCAGTCATCGCCCAAGCTGCGCTATCTCTGGCATCGGGGA 127
Db 2 CGTCCGGCTCTTTTCCCGGCTCAGTCATCGCCCAAGCTGCGCTATCTCTGGCATCGGGGA 61
QY 128 GGAAGAAGCCCGTGCCTTTTCCCGGAGGTTGAAGCGGCATGGAAGAGATTTGCCGAGGA 187
Db 62 GGAAGAAGCCCGTGCCTTTTCCCGGAGGTTGAAGCGGCATGGAAGAGATTTGCCGAGGA 121
QY 188 TGACTGCTGCTCATGATGAGTTGAGCGAAGAACGCGATTTACCATGATGATTTCCGGAAGG 247
Db 122 TGACTGCTGCTCATGATGAGTTGAGCGAAGAACGCGATTTACCATGATGATTTCCGGAAGG 181
QY 248 TGTGGCCATGACGCTTTTAAAGCTGAACTGTTTCTGAGGCGACCTGGGATACAGATTC 307
Db 182 TGTGGCCATGACGCTTTTAAAGCTGAACTGTTTCTGAGGCGACCTGGGATACAGATTC 241
QY 308 GTCCGGGCTTTTCCGACACAGTTCCGGATGTGACCGCGCATGAGCAACCCGAA 367
Db 242 GTCCGGGCTTTTCCGACACAGTTCCGGATGTGACCGCGCATGAGCAACCCGAA 301
QY 368 CAATACCGCGCAGCGCGAAGTCCGCTGCGGTGTCAGATTAATGACAGCGGTGCGGC 427
Db 302 CAATACCGCGCAGCGCGAAGTCCGCTGCGGTGTCAGATTAATGACAGCGGTGCGGC 361
QY 428 GCTGGGATATTACGACGAGGAGCGGTATCTCTGCTGATGCCGACGAAATGACATG 487
Db 362 GCTGGGATATTACGACGAGGAGCGGTATCTCTGCTGATGCCGACGAAATGACATG 421
QY 488 GATACCCCGTGAATTACCCGCGCGCGCCCTCGTTTCATTCACGTTTTTGAACCCGTGA 547
Db 422 GATACCCCGTGAATTACCCGCGCGCGCCCTCGTTTCATTCACGTTTTTGAACCCGTGA 481
QY 548 GGAAGCGGACAGCTCGGGTGCAATGTGTTTACAGCGTGATGGAGCATGAAGATGCT 607
Db 482 GGAAGCGGACAGCTCGGGTGCAATGTGTTTACAGCGTGATGGAGCATGAAGATGCT 541
QY 608 CGACACGCTGCAGAACACCGCAGCT 631
Db 542 CGACACGCTGCAGAACACCGCAGCT 565
```

```
RESULT 12
CD351273/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
```

CD351273 749 bp mRNA linear EST 09-JUL-2003
UI-M-G10-cgh-f-22-0-UI-r1 NIH_BMAP_G10 Mus musculus cDNA clone
IMAGE:6853367 5', mRNA sequence.
CD351273
EST.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 749)
NIH-MGC <http://mgc.nci.nih.gov/>
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Tissue procurement: Dr. Jim Lin, University of Iowa
Email: cgabbs-remail.nih.gov
cDNA Library prepared by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mousefl.html>
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

Seq primer: pYX-5,
Location/Qualifiers
1..749
/organism="Mus musculus"
/mol_type="mRNA"

FEATURES
source

RESULT 13
CR139473/c

LOCUS
DEFINITION

CR139473 733 bp DNA linear GSS 06-JUL-2004
Forward strand read from insert in 3'HPRT insertion targeting and

/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6853367"
/tissue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH BMAP_G10"
/note="Organ: Brain; Vector: pYX-Asc; Site 1: Ecor I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is ACCGAGACAG. This library was created for the University
Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator."

ORIGIN

Query Match 79.7%; Score 563.4; DB 6; Length 749;
Best Local Similarity 99.8%; Pred. No. 9.8e-158;
Matches 564; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 45 ATCTAGCTGTCATCAGATCATATCGTCCGGTCTTTTCCGGCTCAGTCATCGCCCAAGC 104
Db 565 ATCCAGCTGCATCAGGATCATATCGTCCGGTCTTTTCCGGCTCAGTCATCGCCCAAGC 506
QY 105 TGGCGCTATCTCGGCATCGGGGAGGAAGAGCCGTCCTTTCCCGCAGAGTTGAAGCG 164
Db 505 TGGCGCTATCTCGGCATCGGGGAGGAAGAGCCGTCCTTTTCCCGCAGAGTTGAAGCG 446
QY 165 GCATGAAAGAGTTTGGCCGAGGATGACTGCTCTCATTTGACGTTGAGCGAAGAACGACG 224
Db 445 GCATGAAAGAGTTTGGCCGAGGATGACTGCTCTCATTTGACGTTGAGCGAAGAACGACG 386
QY 225 TTTACATGATGATTCGGGAGGTGTGCCATGACGCGCTTTAAACGGTGAACCTGCTT 284
Db 385 TTTACATGATGATTCGGGAGGTGTGCCATGACGCGCTTTAAACGGTGAACCTGCTT 326
QY 285 CAGGCCACTCTGGGATACAGATTCGTCGGCGCTTTTCCGGACACAGTTCCGGATGTCAGC 344
Db 325 CAGGCCACTCTGGGATACAGATTCGTCGGCGCTTTTCCGGACACAGTTCCGGATGTCAGC 266
QY 345 CCGAAGCGCATCAGCAACCCGAAACAATACCGCGCAGACCGGAACTGCCGCGGTGTG 404
Db 265 CCGAAGCGCATCAGCAACCCGAAACAATACCGCGCAGACCGGAACTGCCGCGGTGTG 206
QY 405 CAGATTATGACAGCGGTGCGCGCTGGGATATTAGTCAGGAGACGGGTATCTCTGC 464
Db 205 CAGATTATGACAGCGGTGCGCGCTGGGATATTAGTCAGGAGACGGGTATCTCTGC 146
QY 465 TGGATGCCGAGAAATGGACATGATACCCCTGAGTTACCCGCGCGCGCGCTCTGTC 524
Db 145 TGGATGCCGAGAAATGGACATGATACCCCTGAGTTACCCGCGCGCGCGCTCTGTC 86
QY 525 ATTACAGTTTTTGAACCCCGTGGAGACCGGCGAGACTCGCGGTGCAAAATGTGTTTACAGC 584
Db 85 ATTACAGTTTTTGAACCCCGTGGAGACCGGCGAGACTCGCGGTGCAAAATGTGTTTACAGC 26
QY 585 GTGATGGAGCAGATGAAGATGCTCG 609
Db 25 GTGATGGAGCAGATGAAGATGCTCG 1

ACCESSION	Chromosome engineering clone MHP274b23, genomic survey sequence.
VERSION	CR139473.1 GI:49887345
KEYWORDS	GSS; genome survey sequence; MICER.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS	1. (bases 1 to 733)
	Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L., Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y., Rogers,J. and Bradley,A.
TITLE	Direct Submission
JOURNAL	Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. http://www.sanger.ac.uk/MICER
FEATURES	Location/Qualifiers
source	1..733
	/organism="Mus musculus"
	/mol_type="genomic DNA"
	/db_xref="taxon:10090"
	/clone="MHP274b23"
	/clone_lib="MHP"
ORIGIN	
Query Match	79.5%; Score 562; DB 9; Length 733;
Best Local Similarity	99.3%; Pred. No. 2.6e-157;
Matches	57; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
Qy	56 TCAGGATCATATCGTCGGGCTCTTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGCTATCT 115
Db	733 TCAGGATCATATCGTCGGGCTCTTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGCTATCT 674
Qy	116 GGGCATCGGGAGGAAGAGCCCGTGCTTTTCCGGGAGGTTGAAGCGGCGATGGAAGA 175
Db	673 GGGCATCGGGAGGAAGAGCCCGTGCTTTTCCGGGAGGTTGAAGCGGCGATGGAAGA 614
Qy	176 GTTTGCGGAGGATGACATGCTGCTGCATTGACGTTGAGCGAAACGACGTTTACCATGAT 235
Db	613 GTTTGCGGAGGATGACATGCTGCTGCATTGACGTTGAGCGAAACGACGTTTACCATGAT 554
Qy	236 GATTTCGGGAAGTGTCGCCATGCAAGCTTTAACGGTGAACCTGTTTCAGGCCACCTG 295
Db	553 GATTTCGGGAAGTGTCGCCATGCAAGCTTTAACGGTGAACCTGTTTCAGGCCACCTG 494
Qy	296 GGATACAGTTGCTCGGGCTTTTCCGGACACAGTTCGGATGCTCAGCCGAGCGCAT 355
Db	493 GGATACAGTTGCTCGGGCTTTTCCGGACACAGTTCGGATGCTCAGCCGAGCGCAT 434
Qy	356 CAGCAACCCGAAACAATACCGCGACAGCCGGAACCTGCCG-TGCCGGTGTGCAGATTAATG 414
Db	433 CAGCAACCCGAAACAATACCGCGACAGCCGGAACCTGCCG-TGCCGGTGTGCAGATTAATG 374
Qy	415 ACAGCGTGGCGGCTCGGATATTAAGTCAGCAGGACGCGGTATCTTGGCTGGATGCCG 474
Db	373 ACAGCGTGGCGGCTCGGATATTAAGTCAGCAGGACGCGGTATCTTGGCTGGATGCCG 314
Qy	475 AGAAATGGATCATGATACCCGTCAGTTTACCGGCGGCGCGCTCGTTCTATTCACGTTT 534
Db	313 AGAAATGGATCATGATACCCGTCAGTTTACCGGCGGCGCGCTCGTTCTATTCACGTTT 254
Qy	535 TTGAACCCGTCGAGGACGGGACACTCGCGGTGAAATGTGTTTACAGCGGTGATGGAGC 594
Db	253 TTGAACCCGTCGAGGACGGGACACTCGCGGTGAAATGTGTTTACAGCGGTGATGGAGC 194
Qy	595 AGATGAAGATGCTCGACACGCTGCAGAACACGACGAGT 631
Db	193 AGATGAAGATGCTCGACACGCTGCAGAACACGACGAGT 157

RESULT 14					
CR167678/c					
LOCUS	CR167678	633 bp	DNA	linear	GSS 06-JUN-2004
DEFINITION	Reverse strand read from insert in 3'HPRT insertion targeting and				

chromosome engineering clone MHPPI6lk02, genomic survey sequence.					
ACCESSION	CRI67678				
VERSION	CRI67678.1	GI:49946527			
KEYWORDS	GSS; genome survey sequence; MICER.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
REFERENCE	1. (bases 1 to 633)				
AUTHORS	Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L., Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y., Rogers,J. and Bradley,A.				
TITLE	Direct Submission				
JOURNAL	Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. http://www.sanger.ac.uk/MICER				
FEATURES	Location/Qualifiers				
source	1..633				
	/organism="Mus musculus"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:10090"				
	/clone="MHPPI6lk02"				
	/clone_lib="MHPP"				
ORIGIN					
	Query Match	78.5%;	Score 555;	DB 9;	Length 633;
	Best Local Similarity	100.0%;	Pred. No. 3.2e-155;		
	Matches 555;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	77	TTTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCATCGGGGAGGAAGAAGC	136		
Db	633	TTTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCATCGGGGAGGAAGAAGC	574		
Qy	137	CCGTGCTTTTCCGGGAGGTTGAAGCGCATGGAAGAGTTTGCCGAGATGACTGCTG	196		
Db	573	CCGTGCTTTTCCGGGAGGTTGAAGCGCATGGAAGAGTTTGCCGAGATGACTGCTG	514		
Qy	197	CTGCATTGACGTTGAGCGAAACGCAGTTTACCATGATGATTCGGGAAGGTTGTGGCCAT	256		
Db	513	CTGCATTGACGTTGAGCGAAACGCAGTTTACCATGATGATTCGGGAAGGTTGTGGCCAT	454		
Qy	257	GCACGCCCTTAAACGGTGAACCTGTTTCAGGCCAAGCTGGGATACCAAGTTCGTCGGCGCT	316		
Db	453	GCACGCCCTTAAACGGTGAACCTGTTTCAGGCCAAGCTGGGATACCAAGTTCGTCGGCGCT	394		
Qy	317	TTTTCGGGACACAGTTCGGGATGTCAGCCGAGCGCATCAGCAACCGAACATAACCCG	376		
Db	393	TTTTCGGGACACAGTTCGGGATGTCAGCCGAGCGCATCAGCAACCGAACATAACCCG	334		
Qy	377	CGACAGCCGGAACCTGCGTGTGCGAGTTAATGACAGCGGTGCGCGCTGGGATA	436		
Db	333	CGACAGCCGGAACCTGCGTGTGCGAGTTAATGACAGCGGTGCGCGCTGGGATA	274		
Qy	437	TTTAGTCAGCGAGGACGGGTATCCTCGCTGGATGTCGCGAGAAATGACATGGAATACCCG	496		
Db	273	TTTAGTCAGCGAGGACGGGTATCCTCGCTGGATGTCGCGAGAAATGACATGGAATACCCG	214		
Qy	497	TGAGTTACCCGGCGGGCGGCTCGTTTCATTACAGTTTGTACCCGTCGAGACCGGCA	556		
Db	213	TGAGTTACCCGGCGGGCGGCTCGTTTCATTACAGTTTGTACCCGTCGAGACCGGCA	154		
Qy	557	GACTCGCGGTGCAAAATGTTGTTTTACACGCGTGATGGAGCAGATGAAGATGCTGCACAGCT	616		
Db	153	GACTCGCGGTGCAAAATGTTGTTTTACACGCGTGATGGAGCAGATGAAGATGCTGCACAGCT	94		
Qy	617	GCAGAACACGCACT	631		
Db	93	GCAGAACACGCACT	79		

RESULT 15
CB520716/c
LOCUS CB520716 791 bp mRNA linear EST 09-JUL-2003
DEFINITION UI-M-GIO-cej-j-06-0-UI.r1 NIH RWAP GIO Mus musculus cDNA clone

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 5, 2005, 19:22:55 ; Search time 1994.86 Seconds
(without alignments)
11284.911 Million cell updates/sec

Title: US-10-826-523-41
Perfect score: 3662
Sequence: 1 ctaaatgtaagcgtaata.....atttccccgaaagtgcac 3662

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 5706582 seqs, 3073711274 residues

Total number of hits satisfying chosen parameters: 11413164

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA.*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
- 19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
- 21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3662	100.0	3662	13	US-10-001-189-41
2	2665.4	72.8	3637	15	US-10-211-079-18
3	2665.4	72.8	3637	16	US-10-356-708-32
4	2665.4	72.8	3637	18	US-10-280-913A-32
5	2665.4	72.8	3637	18	US-10-684-134-32
6	2665.4	72.8	3637	18	US-10-637-758-32
7	2664.8	72.8	3637	13	US-10-066-390-4
8	2664.8	72.8	3637	13	US-10-206-030-4
9	2664.8	72.8	3637	15	US-10-211-079-4
10	2664.8	72.8	3637	16	US-10-356-708-17
11	2664.8	72.8	3637	16	US-10-205-772-4
					Sequence 41, Appl
					Sequence 18, Appl
					Sequence 32, Appl
					Sequence 32, Appl
					Sequence 32, Appl
					Sequence 4, Appl
					Sequence 4, Appl
					Sequence 4, Appl
					Sequence 17, Appl
					Sequence 4, Appl

C 12	2664.8	72.8	3637	18	US-10-280-913A-17	Sequence 17, Appl
C 13	2664.8	72.8	3637	18	US-10-684-134-17	Sequence 17, Appl
C 14	2664.8	72.8	3637	18	US-10-637-758-17	Sequence 17, Appl
C 15	2656.4	72.5	3637	13	US-10-066-390-3	Sequence 3, Appl
C 16	2656.4	72.5	3637	13	US-10-206-030-3	Sequence 3, Appl
C 17	2656.4	72.5	3637	15	US-10-211-079-3	Sequence 3, Appl
C 18	2656.4	72.5	3637	16	US-10-356-708-31	Sequence 31, Appl
C 19	2656.4	72.5	3637	16	US-10-205-772-3	Sequence 3, Appl
C 20	2656.4	72.5	3637	18	US-10-280-913A-31	Sequence 31, Appl
C 21	2656.4	72.5	3637	18	US-10-684-134-31	Sequence 31, Appl
C 22	2656.4	72.5	3637	18	US-10-637-758-31	Sequence 31, Appl
C 23	2609.4	71.3	4001	11	US-09-860-763-13	Sequence 13, Appl
C 24	2484.6	67.8	3774	11	US-09-860-763-9	Sequence 9, Appl
C 25	2469.4	67.4	5277	16	US-10-412-872-16	Sequence 16, Appl
C 26	2464.6	67.3	3908	17	US-10-014-099F-70	Sequence 70, Appl
C 27	2451.8	67.0	3928	15	US-10-128-590-45	Sequence 45, Appl
C 28	2451.8	67.0	3928	16	US-10-128-587A-45	Sequence 45, Appl
C 29	2451.8	67.0	3928	17	US-10-128-578B-45	Sequence 45, Appl
C 30	2446.8	66.8	3715	11	US-09-860-763-11	Sequence 11, Appl
C 31	2446.8	66.8	3927	17	US-10-014-099F-71	Sequence 71, Appl
C 32	2271.2	62.2	4119	9	US-09-993-170-1	Sequence 1, Appl
C 33	2271.2	62.0	4059	18	US-10-655-702-1	Sequence 1, Appl
C 34	2230.6	60.9	5250	18	US-10-474-070-51	Sequence 51, Appl
C 35	2216.2	60.5	4205	10	US-09-883-848A-48	Sequence 48, Appl
C 36	2216.2	60.5	4205	19	US-10-652-686-48	Sequence 48, Appl
C 37	2213	60.4	4205	10	US-09-883-848A-35	Sequence 35, Appl
C 38	2213	60.4	4205	19	US-10-652-686-35	Sequence 35, Appl
C 39	2172.8	59.3	7803	17	US-10-014-099F-82	Sequence 82, Appl
C 40	2172.8	59.3	8167	17	US-10-014-099F-83	Sequence 83, Appl
C 41	2172.6	59.3	10600	16	US-10-356-708-1	Sequence 1, Appl
C 42	2172.6	59.3	10600	18	US-10-280-913A-1	Sequence 1, Appl
C 43	2172.6	59.3	10600	18	US-10-684-134-1	Sequence 1, Appl
C 44	2172.6	59.3	10600	18	US-10-637-758-1	Sequence 1, Appl
C 45	2172.6	59.3	10624	16	US-10-356-708-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1

US-10-001-189-41
; Sequence 41, Application US/10001189
; Publication No. US20020173634A1
; GENERAL INFORMATION:
; APPLICANT: FRASER JR., MALCOLM J.
; APPLICANT: LI, XU
; APPLICANT: BEAM, TERESA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
; FILE REFERENCE: 835910-92098
; CURRENT APPLICATION NUMBER: US/10/001,189
; CURRENT FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/244,984
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/244,677
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 3662
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: pXL-Bac
; OTHER INFORMATION: sequence
US-10-001-189-41

Query Match 100.0%; Score 3662; DB 13; Length 3662;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 3662; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTAATGTAAGCGTTAAATTTTGTGTAATAATTCGCGTTAAATTTTGTGTAATTCAGCTC 60

[illegible]

1081	Db	CAAGCTGGCGCTATCTTGGGCATCGGGAGGAAGAGCCCGTGCCTTTTCCGCGAGGTTG	1141
1141	Qy	AAGCGGCATCGAAAGAGTTTTCGCGAGGATGACTGCTGCTGCATTGACGTTTGAGCGAAAC	1200
1141	Db	AAGCGGCATCGAAAGAGTTTTCGCGAGGATGACTGCTGCTGCATTGACGTTTGAGCGAAAC	1200
1201	Qy	GCACGTTTACCATGATGATTCGGGAAGTGTGCCATGCAAGCCTTTAAACGGTGAACTGT	1260
1201	Db	GCACGTTTACCATGATGATTCGGGAAGTGTGCCATGCAAGCCTTTAAACGGTGAACTGT	1260
1261	Qy	TGCTTCAGGCCACTGGGATACCAAGTTCGTCGGGCTTTTCCGACACAGTTCGGATGG	1320
1261	Db	TGCTTCAGGCCACTGGGATACCAAGTTCGTCGGGCTTTTCCGACACAGTTCGGATGG	1320
1321	Qy	TCAGCCGAAGCGCATCAGCAACCCGAAACAATACCGCGCACAGCCGGAACCTGCGTGC	1380
1321	Db	TCAGCCGAAGCGCATCAGCAACCCGAAACAATACCGCGCACAGCCGGAACCTGCGTGC	1380
1381	Qy	GTGTGCAGATTAAATGACAGCGGTGCGCGCTGGGATATTACGTACGCGAGGA	1440
1381	Db	GTGTGCAGATTAAATGACAGCGGTGCGCGCTGGGATATTACGTACGCGAGGA	1440
1441	Qy	CTGGCTGGATGCGCAGAAATGACATGATACCCCGTGAGTTTACCCGCGGGCGCGCTT	1500
1441	Db	CTGGCTGGATGCGCAGAAATGACATGATACCCCGTGAGTTTACCCGCGGGCGCGCTT	1500
1501	Qy	GGCGTAATCATGTCTATAGCTGTTTCTCTGTGCAATTTGTTATCCGCTCACAATTCACA	1560
1501	Db	GGCGTAATCATGTCTATAGCTGTTTCTCTGTGCAATTTGTTATCCGCTCACAATTCACA	1560
1561	Qy	CAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTTGGGTCCTTAATGAGTGAGCTAACT	1620
1561	Db	CAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTTGGGTCCTTAATGAGTGAGCTAACT	1620
1621	Qy	CACATTAATTGCGTTGGCTCACTGCGCGCTTTCCAGTCGGGAAAACCTGTGTCGACGCT	1680
1621	Db	CACATTAATTGCGTTGGCTCACTGCGCGCTTTCCAGTCGGGAAAACCTGTGTCGACGCT	1680
1681	Qy	GCATTAATGAATCGGCCAACCGCGGGGAGAGCGGTTTCGCTATTGGGCGCTCTTCCGC	1740
1681	Db	GCATTAATGAATCGGCCAACCGCGGGGAGAGCGGTTTCGCTATTGGGCGCTCTTCCGC	1740
1741	Qy	TTCTCTCGCTCACTGACTCGCTGCGCTCGGTGCTGCGCTGCGCGAGCGGTATCAGCTCA	1800
1741	Db	TTCTCTCGCTCACTGACTCGCTGCGCTCGGTGCTGCGCTGCGCGAGCGGTATCAGCTCA	1800
1801	Qy	CTCAAAGCGGTAAATACGGTTATCCACAGAAATCAGGGGATAACGCGAGGAAGAACATGTC	1860
1801	Db	CTCAAAGCGGTAAATACGGTTATCCACAGAAATCAGGGGATAACGCGAGGAAGAACATGTC	1860
1861	Qy	AGCAAAAGGCCACAAAGGCCAGGAACCGTAAAAAGGCGCGTGTGTCGCGCTTTTCCCA	1920
1861	Db	AGCAAAAGGCCACAAAGGCCAGGAACCGTAAAAAGGCGCGTGTGTCGCGCTTTTCCCA	1920
1921	Qy	TAGGCTCCGCCCTCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGAAA	1980
1921	Db	TAGGCTCCGCCCTCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGAAA	1980
1981	Qy	CCGACAGGACTATAAAGATACAGCGGTTTCCCTTGGAAGCTCCCTGTGCGCTCTCC	2040
1981	Db	CCGACAGGACTATAAAGATACAGCGGTTTCCCTTGGAAGCTCCCTGTGCGCTCTCC	2040
2041	Qy	TGTTCCGACCTCGCGCTTACCGGATACCTGTCCGCGCTTCTCCCTTCGGGAAGCGTGGC	2100
2041	Db	TGTTCCGACCTCGCGCTTACCGGATACCTGTCCGCGCTTCTCCCTTCGGGAAGCGTGGC	2100
2101	Qy	GCTTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTTCGCTCCAGCT	2160
2101	Db	GCTTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTTCGCTCCAGCT	2160
2161	Qy	GGGCTGTGTGCAAGAACCCCGCTTACGCGCGACCGCTGCGCTTATTCGGTAACTATCG	2220
2161	Db	GGGCTGTGTGCAAGAACCCCGCTTACGCGCGACCGCTGCGCTTATTCGGTAACTATCG	2220


```
QY 2221 TCTTGAGTCCAAACCCGGTAAAGACACGACCTTATCGCCACTGGCAGCAGCCACTGGTAAACAG 2280
Db 2221 TCTTGAGTCCAAACCCGGTAAAGACACGACCTTATCGCCACTGGCAGCAGCCACTGGTAAACAG 2280
QY 2281 GATTAGCAGCAGAGGTATGTAGGCGGTGCTACAGAGTTCTTCAGTGGTGGCCCTAACTA 2340
Db 2281 GATTAGCAGCAGAGGTATGTAGGCGGTGCTACAGAGTTCTTCAGTGGTGGCCCTAACTA 2340
QY 2341 CGGCTACACTAGAAGGACAGTATTTGGGTATCTGCGCTCTGCTGAAAGCCAGTTACCTTCGG 2400
Db 2341 CGGCTACACTAGAAGGACAGTATTTGGGTATCTGCGCTCTGCTGAAAGCCAGTTACCTTCGG 2400
QY 2401 AAAAGAGTTGGTAGTCTTGATTCGGGCAAAACAAACACCGCTGGTAGCGGTGGTTTTTT 2460
Db 2401 AAAAGAGTTGGTAGTCTTGATTCGGGCAAAACAAACACCGCTGGTAGCGGTGGTTTTTT 2460
QY 2461 TGTGTCAGACGACAGATTACGGCGAGAAAAGGATCTCAAGAGATCCCTTTCATCTT 2520
Db 2461 TGTGTCAGACGACAGATTACGGCGAGAAAAGGATCTCAAGAGATCCCTTTCATCTT 2520
QY 2521 TTCTACGGGGTCTGACGCTCAGTGGAAACGAAACTCAGCTTAAGGATTTTGGTCATGAG 2580
Db 2521 TTCTACGGGGTCTGACGCTCAGTGGAAACGAAACTCAGCTTAAGGATTTTGGTCATGAG 2580
QY 2581 ATTATCAAAAAGGATCTTCACTAGATCCTTTTAAATTAATAAATGAAGTTTAAATCAAT 2640
Db 2581 ATTATCAAAAAGGATCTTCACTAGATCCTTTTAAATTAATAAATGAAGTTTAAATCAAT 2640
QY 2641 CTAAGATATATAGTAAACTTGTCTGACAGTTACCAATGCTTAATCAGTAGGACAC 2700
Db 2641 CTAAGATATATAGTAAACTTGTCTGACAGTTACCAATGCTTAATCAGTAGGACAC 2700
QY 2701 TATCTCAGGATCTGTCTATTTTCGTTTCATCTCATAGTTGCTGACTCCCGTCTGTAGAT 2760
Db 2701 TATCTCAGGATCTGTCTATTTTCGTTTCATCTCATAGTTGCTGACTCCCGTCTGTAGAT 2760
QY 2761 AACTACGATACGGGAGGGCTTACCATTCTGGCCCCAGTGTGCAATGATACCCGAGACCC 2820
Db 2761 AACTACGATACGGGAGGGCTTACCATTCTGGCCCCAGTGTGCAATGATACCCGAGACCC 2820
QY 2821 AGCTCACCGGCTCCAGATTTATCAGCAATAACAGCAGCCAGCGGAGGCGGAGCGAG 2880
Db 2821 AGCTCACCGGCTCCAGATTTATCAGCAATAACAGCAGCCAGCGGAGGCGGAGCGAG 2880
QY 2881 AAGTGGTCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTCGGCGAAGCTAG 2940
Db 2881 AAGTGGTCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTCGGCGAAGCTAG 2940
QY 2941 AGTAAGTAGTTCCGCAATTAAGTTTGGCAGAAAGTTGTGCAATTTGCTACAGGCATCGT 3000
Db 2941 AGTAAGTAGTTCCGCAATTAAGTTTGGCAGAAAGTTGTGCAATTTGCTACAGGCATCGT 3000
QY 3001 GGTGTACGCTCGTGTGGTATGAGCTTATTCAGCTCCGCTCCCAACGATCAAGGGG 3060
Db 3001 GGTGTACGCTCGTGTGGTATGAGCTTATTCAGCTCCGCTCCCAACGATCAAGGGG 3060
QY 3061 AGTTACATGATCCCGATGTTGCAAAAAGCGTTAGCTCTTCGCTCCGATCGT 3120
Db 3061 AGTTACATGATCCCGATGTTGCAAAAAGCGTTAGCTCTTCGCTCCGATCGT 3120
QY 3121 TGTGAGAAGTAAGTTGGCCGAGTGTATCACTCATGTTTATGGCAGCACTGCATAATTC 3180
Db 3121 TGTGAGAAGTAAGTTGGCCGAGTGTATCACTCATGTTTATGGCAGCACTGCATAATTC 3180
QY 3181 TCTTACTGTATGCCATCCGTAAGATGCTTTTCTGTGATGCTGTGATGACTCAACCAAGTC 3240
Db 3181 TCTTACTGTATGCCATCCGTAAGATGCTTTTCTGTGATGCTGTGATGACTCAACCAAGTC 3240
QY 3241 ATTCTGAGAAATAGTATCGGGGACCGAGTTGCTTCGCCGGGTCAATACGGGATAA 3300
Db 3241 ATTCTGAGAAATAGTATCGGGGACCGAGTTGCTTCGCCGGGTCAATACGGGATAA 3300
```

```
QY 3301 TACCGGGCCACATAGCAGAACTTTAAAGTGTCTCATCTTGGAAAACGTTCTTCGGGGCG 3360
Db 3301 TACCGGGCCACATAGCAGAACTTTAAAGTGTCTCATCTTGGAAAACGTTCTTCGGGGCG 3360
QY 3361 AAAACTCTCAAGGATCTTACCCTGTTGAGATCCAGTTCGATGAACCACTCGTGACCC 3420
Db 3361 AAAACTCTCAAGGATCTTACCCTGTTGAGATCCAGTTCGATGAACCACTCGTGACCC 3420
QY 3421 CAACTGATCTTCAGCATCTTTTACTTTTCCAGCGTTTCTGGGTGAGCAAAAACAGGAAG 3480
Db 3421 CAACTGATCTTCAGCATCTTTTACTTTTCCAGCGTTTCTGGGTGAGCAAAAACAGGAAG 3480
QY 3481 GCAAAATCGCCCAAAAAGGGAATAAGCGGCAGACGGAATGTTGAATCTCATCTCTT 3540
Db 3481 GCAAAATCGCCCAAAAAGGGAATAAGCGGCAGACGGAATGTTGAATCTCATCTCTT 3540
QY 3541 CTTTTTCAATATATTGAAGCATTTTATCAGGGTTATCTCTCATGCGGATACATATT 3600
Db 3541 CTTTTTCAATATATTGAAGCATTTTATCAGGGTTATCTCTCATGCGGATACATATT 3600
QY 3601 TGAATGTATTTAGAAAAATAAACAATAGGGGTTCCGCGCACATTTCCCGAAAAAGTGCC 3660
Db 3601 TGAATGTATTTAGAAAAATAAACAATAGGGGTTCCGCGCACATTTCCCGAAAAAGTGCC 3660
QY 3661 AC 3662
Db 3661 AC 3662
```

RESULT 2

```
US-10-211-079-18/c
; Sequence 18, Application US/10211079
; Publication No. US20030148315A1
; GENERAL INFORMATION:
; APPLICANT: Padgett, Hal S.
; APPLICANT: Vaehongs, Andrew A.
; APPLICANT: Vojdani, Fakhrish S.
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding CEL I
; TITLE OF INVENTION: Endonuclease and Methods of Use Thereof
; FILE REFERENCE: P-LG 5381
; CURRENT APPLICATION NUMBER: US/10/211,079
; CURRENT FILING DATE: 2002-08-01
; PRIOR APPLICATION NUMBER: US 10/098,155
; PRIOR FILING DATE: 2002-03-14
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 3637
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-10-211-079-18
```

```
Query Match 72.8%; Score 2665.4; DB 15; Length 3637;
Best Local Similarity 83.7%; Pred. No. 0;
Matches 3065; Conservative 0; Mismatches 571; Indels 27; Gaps 3;
```

```
QY 1 CTAATTTGTAAGCGTTAATATTTTGTAAAAATTCGGTTTAAATTTTGTAAATCAGCTC 60
Db 3637 CTAATTTGTAAGCGTTAATATTTTGTAAAAATTCGGTTTAAATTTTGTAAATCAGCTC 3578
QY 61 ATTTTAAACCAATAGGCGGAATCGCAAAATCCCTTATAAATCAAAAGAAATAGACCGA 120
Db 3577 ATTTTAAACCAATAGGCGGAATCGCAAAATCCCTTATAAATCAAAAGAAATAGACCGA 3518
QY 121 GATAGGGTTGAGTGTGTTTCCAGTTTGGAAACAGAGTCCACTATTAAAGAACGTGGACTC 180
Db 3517 GATAGGGTTGAGTGTGTTTCCAGTTTGGAAACAGAGTCCACTATTAAAGAACGTGGACTC 3458
QY 181 CAACTCAAGGGCGGAAAAACCGCTCTATCAGGGCGAGTGGCCCACTACGTGAACCATCACC 240
Db 181 CAACTCAAGGGCGGAAAAACCGCTCTATCAGGGCGAGTGGCCCACTACGTGAACCATCACC 240
```


Db 3457 CAACGCTCAAGGGCGAAACCCGTCATATCAGGGCGATGGCCCACTACGTGAACCATCAC 3398
Qy 241 CTAATCAAGTTTTTTGGGGTCGAGGTGCGTAAACACATAAATCGGAACCCCTAAAGGGAG 300
Db 3397 CTAATCAAGTTTTTTGGGGTCGAGGTGCGTAAACACATAAATCGGAACCCCTAAAGGGAG 3338
Qy 301 CCCCGATTATAGACTTGA CGGGGAAAGCGCGCAACCGTGGCGAGAAAGGAAGGAAGAA 360
Db 3337 CCCCCGATTAGACTTGA CGGGGAAAGCGCGCAACCGTGGCGAGAAAGGAAGGAAGAA 3278
Qy 361 AGCGAAAGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGTACGCTGCGCGTAACAC 420
Db 3277 AGCGAAAGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGTACGCTGCGCGTAACAC 3218
Qy 421 CACACCCCGCGCTTAATGCGCGCTACAGGGCGCTCCCATTCGCCATTTCAGGCTCG 480
Db 3217 CACACCCCGCGCTTAATGCGCGCTACAGGGCGCTCCCATTCGCCATTTCAGGCTCG 3158
Qy 481 CAACCTGTTGGGAAGGGCGATCGGTGCGGGCCCTCTTCGCTATTACGCCAGCTGGCGAAAG 540
Db 3157 CAACCTGTTGGGAAGGGCGATCGGTGCGGGCCCTCTTCGCTATTACGCCAGCTGGCGAAAG 3098
Qy 541 GGGATGCTGCGAAGCGGCAATTAAGTTGGGTAAAGCCAGGGTTTCCCAAGTCACGAGCTG 600
Db 3097 GGGATGCTGCGAAGCGGCAATTAAGTTGGGTAAAGCCAGGGTTTCCCAAGTCACGAGCTG 3038
Qy 601 TAAACGACGGCCAGTGAGCGGCTCGTTCAATTCACCGTTTTTGAACCCGTTGGAGGACGG 660
Db 3037 TAAACGACGGCCAGTGAGCGGCGCTTAATGCACTATAGGGCGAATTGGAGCTCCA 2978
Qy 661 GCAGACTCGGGTGCAAAATGTTTTACAGCGTGTATGGAGCAGATGAAGATGCTCGACAC 720
Db 2977 CCGCGTGGCGGCGCTTAGNACTAGTGTGATGCCCGGCGTCAGGAATCTTATTGTT 2918
Qy 721 GCTGCAGAACCGCAGCTAGATTAACCTCAGAAAGATTAATCATATTGTGACGTAGTTAA 780
Db 2917 ATAGTTTATCCATGTCATGTTAATCCAGCAGCAGTTTACAACTCAAGAGGACCATGT 2858
Qy 781 AGATAATC-ATGCGTAAATTCAGCATGGATCTGTAATACGACTCACTATAGGGCGAA 839
Db 2857 GGTACGCTTTTCGTTGGGATCTTTTCGAAAGGGCAGATTGTGTCACAGGTAATGGTTGT 2798
Qy 840 TTGGGTACCGGGCCCCCTCGAGGTGACGCGTATCGATAAGCTTTGATATCGAATTCCTG 899
Db 2797 CTGGTAAAGGACAGGGCCATCGCCAAATTGGAGTATTTTGTGATAATGGTCTGCTAGT 2738
Qy 900 CAGCCCGGGGATCCATAGTTCTAGAGGGCGCCACCGGGTGGAGCTCAGCTTTTG 959
Db 2737 GAAACGGATCCATCTTTCAATGTTGTGGCGAATTTTGAAGTTAGCTTTGATTCATTTCTTT 2678
Qy 960 TTCCCTTTAGTGAGGGTTAATTAGATCCCATGCGTCAATTTTACGACACTATCTTTCTA 1019
Db 2677 GTTTGCTGCGGTGATGATATCATTTGTGTAGTTTAAAGTTGTACTCGAGTTTGTGTCGA 2618
Qy 1020 GGGTTAATCTAGCTGATCAGGATCATATCGTGGGTCTTTTTTCGGCTCAGTCATCGC 1079
Db 2617 GAAATGTTCCATCTCTTTTAAATCAATACCTTTTAACTCGATAC- - - -GATTAAACAG 2563
Qy 1080 CCAAGCTGGGTATCTGGGCAATCGGGAGGAAGAGCCCGTGGCTTTTCCCGCGAGGTT 1139
Db 2562 GGTATCACCTTCAAACTTTGACTTTCAGCACGGCTCTTTGATTTCCCGTCACTTTTGAAGA 2503
Qy 1140 GAAGCGGCAATGAAGAGTTTGGCGAGGATGACTGCTGCTGCAATTGAGCTTTCAGCGAAGA 1199
Db 2502 TATAGTGGCTTCTGTA CATTAACCTTTCGGGATGGCACTCTTGAAAGAGTATGCGCGTTT 2443
Qy 1200 CGCAGTTTACCATGATGATTTCGGGAAGGTGTGGCCATGCGCCCTTTTAAAGGTTGAAC 1259
Db 2442 CATATGATCCGGATAACGAGAAAGCAATTGAACCATGAGAGAAAGTAGTGACAAGTGT 2383
Qy 1260 TTCGTTACGGCCACTCGGGAATACAGTTTCGTCGGGCTTTTCCGAGACACAGTTCCGGATG 1319
Db 2382 TGGCCATGGAACAGGTAGTTTTTCCAGTAGTGTGCAATAAAATTTAAGTGTGAAGCTTTCCGTA 2323

Qy 1320 GTCAGCCCGAAGCGCATCAGCAACCCGAAACAATACCGGCGACAGCCGGAATGCGCGTGCC 1379
Db 2322 TGTAGCATCACCTTCCACCTCTCCACTGACAGAAAAATTTGTGCCCATTAACATCACCATC 2263
Qy 1380 GGTGTGCAAGTAAATGACAGCGGTGCGGCGTGGGATATTAAGTCAGCAGGAGACGGGTAT 1439
Db 2262 TAAATCAACAAGAAAT- - - - -TGGGACAACTCCAGTGAAGATTTC 2224
Qy 1440 CCTGGCTGGATGCCGAGAAATGGACATGGATACCCCGTGAGTTACCCGGGGCGCGCT 1499
Db 2223 TTCTCTTTTACTCATPCGATCCAGCTTTTGTTCCTTTTAGTGAGGGTTAATTGTCGCGCT 2164
Qy 1500 TGGCGTAAATCATGGTCATAGCTGTTTCTGTGTGAATTTGTTATCCGCTCAAAATTCAC 1559
Db 2163 TGGCGTAAATCATGGTCATAGCTGTTTCTGTGTGAATTTGTTATCCGCTCAAAATTCAC 2104
Qy 1560 ACAACATACGAGCGGAAAGATAAAGTGTAAAGCTCGGGTGCCTTAATGAGTGAGCTAAC 1619
Db 2103 ACAACATACGAGCGGAAAGATAAAGTGTAAAGCTCGGGTGCCTTAATGAGTGAGCTAAC 2044
Qy 1620 TCACATTAATTTGCGTTGCGCTCACTGCCCCGCTTTCCAGTCGGGAAACCTGTCGTCACAG 1679
Db 2043 TCACATTAATTTGCGTTGCGCTCACTGCCCCGCTTTCCAGTCGGGAAACCTGTCGTCACAG 1984
Qy 1680 TGCATTTAATGAATCGGCCAAACCGCGGGGAGAGCGGTTTTCGCTATTGGGGCTCTTCCG 1739
Db 1983 TGCATTTAATGAATCGGCCAAACCGCGGGGAGAGCGGTTTTCGCTATTGGGGCTCTTCCG 1924
Qy 1740 CTTCTCGCTCACTGACTGCTGCGCTCGGTCGGTTCGGTTCGGCGAGCGGTATCAGCTC 1799
Db 1923 CTTCTCGCTCACTGACTGCTGCGCTCGGTCGGTTCGGTTCGGCGAGCGGTATCAGCTC 1864
Qy 1800 ACTCAAGGCGGTAATACGTTTATCCACAGAAATCAGGGGATAACGACAGGAAGACATGT 1859
Db 1863 ACTCAAGGCGGTAATACGTTTATCCACAGAAATCAGGGGATAACGACAGGAAGACATGT 1804
Qy 1860 GAGCAAAAGGCGCAGCAAAAGGCGCAGGAACCGTAAAGGCGCGCTTGTGCGGTTTTTCC 1919
Db 1803 GAGCAAAAGGCGCAGCAAAAGGCGCAGGAACCGTAAAGGCGCGCTTGTGCGGTTTTTCC 1744
Qy 1920 ATAGGCTCGGCGGCTGACAGGATCAAAATAATCGACGCTCAAGTCAGAGGTGGCGAA 1979
Db 1743 ATAGGCTCGGCGGCTGACGAGCATCAAAATAATCGACGCTCAAGTCAGAGGTGGCGAA 1684
Qy 1980 ACCGACAGGACTAAAGATACAGGCGTTTCCCGCTGGAAGCTCCCTGTCGCTCTC 2039
Db 1683 ACCGACAGGACTAAAGATACAGGCGTTTCCCGCTGGAAGCTCCCTGTCGCTCTC 1624
Qy 2040 CTGTTCCGACCTCGCGCTTACCAGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGG 2099
Db 1623 CTGTTCCGACCTCGCGCTTACCAGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGG 1564
Qy 2100 CGTTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCCGTTGAGTGTAGTCTGCTCAAGC 2159
Db 1563 CGTTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCCGTTGAGTGTAGTCTGCTCAAGC 1504
Qy 2160 TGGGCTGTGTCAGCAAGAACCCCGTTACGCGCGAGCGCTGCGCTTATCCGCTAACTATC 2219
Db 1503 TGGGCTGTGTCAGCAAGAACCCCGTTACGCGCGAGCGCTGCGCTTATCCGCTAACTATC 1444
Qy 2220 GTCTTGAGTCCAAACCCGGAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACA 2279
Db 1443 GTCTTGAGTCCAAACCCGGAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACA 1384
Qy 2280 GGATTAGCAGCAGGATGTAGCGGTGTCTACAGATTTCTTGAAGTGTGGGCTTAAC 2339
Db 1383 GGATTAGCAGCAGGATGTAGCGGTGTCTACAGATTTCTTGAAGTGTGGGCTTAAC 1324
Qy 2340 ACGGCTACCTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCAGTTACTTCG 2399
Db 1323 ACGGCTACCTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCAGTTACTTCG 1264

QY 2400 GAAAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACACCCGCTGGTAGCGGTGTTTTT 2459
DB 1263 GAAAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACCCGCTGGTAGCGGTGTTTTT 1204
QY 2460 TTGTTTGCAGCAGCAGATAGCGCAGAAAAAGGATCTCAAGAAAGATCTTTTGTATCT 2519
DB 1203 TTGTTTGCAGCAGCAGATAGCGCAGAAAAAGGATCTCAAGAAAGATCTTTTGTATCT 1144
QY 2520 TTCTACCGGGCTGACGCTCAGTGGAACGAAACCTCACGTTAAAGGATTTTGGTCATGA 2579
DB 1143 TTCTACCGGGCTGACGCTCAGTGGAACGAAACCTCACGTTAAAGGATTTTGGTCATGA 1084
QY 2580 GATTATCAAAAAGGATCTTACCTAGATCTTTTAAATTTAAATTTAAATTTAAATCAA 2639
DB 1083 GATTATCAAAAAGGATCTTACCTAGATCTTTTAAATTTAAATTTAAATTTAAATCAA 1024
QY 2640 TCTAAAGTATATAGTAAATCTTGCTCTGACAGTTTACCAATGCTTAAATCAGTAGGCAC 2699
DB 1023 TCTAAAGTATATAGTAAATCTTGCTCTGACAGTTTACCAATGCTTAAATCAGTAGGCAC 964
QY 2700 CTATCTCAGCGATCTGTCTATTTCGTTTCATCCATAGTTCCCTGACTCCCGTGTGTAGA 2759
DB 963 CTATCTCAGCGATCTGTCTATTTCGTTTCATCCATAGTTCCCTGACTCCCGTGTGTAGA 904
QY 2760 TAACTACGATACGGAGGGCTTACCATCTGCGCCGAGTCTGCAATGATACCGCGAGACC 2819
DB 903 TAACTACGATACGGAGGGCTTACCATCTGCGCCGAGTCTGCAATGATACCGCGAGACC 844
QY 2820 CACGCTCACCGCTCCAGATTTATCAGCAATTAACAGCAGCGAGGAGGGCGGAGCGCA 2879
DB 843 CACGCTCACCGCTCCAGATTTATCAGCAATTAACAGCAGCGAGGAGGGCGGAGCGCA 784
QY 2880 GAAGTGGTCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTCGGGAGCTTA 2939
DB 783 GAAGTGGTCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTCGGGAGCTTA 724
QY 2940 GAGTAAGTAGTTCCGAGTTAATAGTTTCCGCAACGTTGTCGATGCTACAGCATCG 2999
DB 723 GAGTAAGTAGTTCCGAGTTAATAGTTTCCGCAACGTTGTCGATGCTACAGCATCG 664
QY 3000 TGGTGTACGCTCGTCTGTTTGGTATGGCTTCATTGAGCTCCGTTCCCAACGATCAAGC 3059
DB 663 TGGTGTACGCTCGTCTGTTTGGTATGGCTTCATTGAGCTCCGTTCCCAACGATCAAGC 604
QY 3060 GAGTTACATGATCCCCCATGTTGTGCAAAAAGCGTTAGCTCTTCCGTCCTCCGATCG 3119
DB 603 GAGTTACATGATCCCCCATGTTGTGCAAAAAGCGTTAGCTCTTCCGTCCTCCGATCG 544
QY 3120 TTGTGAGAAAGTAAAGTTGGCCGAGTGTATCACTCATGTTATGGCAGCACTGCATAATT 3179
DB 543 TTGTGAGAAAGTAAAGTTGGCCGAGTGTATCACTCATGTTATGGCAGCACTGCATAATT 484
QY 3180 CTCCTTACTGTCATGCCATCCGTAAGATGCTTTCTGTGACTGTGAGTACTCAACCAAGT 3239
DB 483 CTCCTTACTGTCATGCCATCCGTAAGATGCTTTCTGTGACTGTGAGTACTCAACCAAGT 424
QY 3240 CATTTCTGAGATAGTGTATGCGCGCAGCCAGTTGCTCTTTCGCGCGGTCAATACGGGATA 3299
DB 423 CATTTCTGAGATAGTGTATGCGCGCAGCCAGTTGCTCTTTCGCGCGGTCAATACGGGATA 364
QY 3300 ATACCGCGCCACATAGCAGAACTTTTAAAGTGTCTCATCTTGGTGAACCGTTCTTTCGGGCG 3359
DB 363 ATACCGCGCCACATAGCAGAACTTTTAAAGTGTCTCATCTATGGAACCGTTCTTTCGGGCG 304
QY 3360 GAAACTCTCAGGATCTTACCGCTGTTCAGATCCAGTTTCGATGTAACCACTCTGTCAC 3419
DB 303 GAAACTCTCAGGATCTTACCGCTGTTCAGATCCAGTTTCGATGTAACCACTCTGTCAC 244
QY 3420 CCAACTGATCTTCAGCATCTTTTACTTTTCAACAGCGTTTCTGGTGTAGCAAAAACAGGAA 3479
DB 243 CCAACTGATCTTCAGCATCTTTTACTTTTCAACAGCGTTTCTGGTGTAGCAAAAACAGGAA 184
QY 3480 GGCAAAATCCGCAAAAAGGGAATAAGGGCGACACGGAAATGTTGAAATACTCATCTCT 3539

DB 183 GGCAAAATCCGCAAAAAGGGAATAAGGGCGACACGGAAATGTTGAAATCTCATCTCT 124
QY 3540 TCCTTTTCAATATTATTGAAGCAATTTATCAGGGTTATTGTCTCATGAGCGGATACATAT 3599
DB 123 TCCTTTTCAATATTATTGAAGCAATTTATCAGGGTTATTGTCTCATGAGCGGATACATAT 64
QY 3600 TTGAATGTATTAGAAAAATAAAATAGGGGTTCCGGCGCACATTTCCCGAAAAAGTGC 3659
DB 63 TTGAATGTATTAGAAAAATAAAATAGGGGTTCCGGCGCACATTTCCCGAAAAAGTGC 4
QY 3660 CAC 3662
DB 3 CAC 1
RESULT 3
US-10-356-708-32/c
; Sequence 32, Application US/10356708
; Publication No. US20030157682A1
; GENERAL INFORMATION:
; APPLICANT: Large Scale Biology Corporation
; TITLE OF INVENTION: MISMATCH ENDONUCLEASES AND METHODS OF USE THEREOF
; FILE REFERENCE: P-LG 10100
; CURRENT APPLICATION NUMBER: US/10/356,708
; CURRENT FILING DATE: 2003-01-31
; PRIOR APPLICATION NUMBER: 60/353,722
; PRIOR FILING DATE: 2000-02-01
; PRIOR APPLICATION NUMBER: 10/098,155
; PRIOR FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 10/211,079
; PRIOR FILING DATE: 2002-08-01
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32
; LENGTH: 3637
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Encodes Cycle 3 BFP gene
US-10-356-708-32
Query Match 72.8%; Score 2665.4; DB 16; Length 3637;
Best Local Similarity 83.7%; Pred. No. 0;
Matches 3065; Conservative 0; Mismatches 571; Indels 27; Gaps 3;
QY 1 CTAATTTGTAAGCGTTAATTTTGTAAAAATTCGCGTTAAATTTTGTAAATCAGCTC 60
DB 3637 CTAATTTGTAAGCGTTAATTTTGTAAAAATTCGCGTTAAATTTTGTAAATCAGCTC 3578
QY 61 ATTTTAAACCAATAGCGCAAAATCGCAAAATCCCTTATAATCAAAAAGAAATAGACCGA 120
DB 3577 ATTTTAAACCAATAGCGCAAAATCGCAAAATCCCTTATAATCAAAAAGAAATAGACCGA 3518
QY 121 GATAGGTTGAGTGTGTTCCAGTTTGGAAACAGAGTCCACTTATTAAGAAACGTGACTC 180
DB 3517 GATAGGTTGAGTGTGTTCCAGTTTGGAAACAGAGTCCACTTATTAAGAAACGTGACTC 3458
QY 181 CAACGTCAAAGCGCGAAACCGTCTATCAGGGCGATGCGCCACTACGTGAACCATCACC 240
DB 3457 CAACGTCAAAGCGCGAAACCGTCTATCAGGGCGATGCGCCACTACGTGAACCATCACC 3398
QY 241 CTAATCAAGTTTTTTGGGTCGAGGTGCGTAAAGCACTAAATCGGAACCTTAAAGGAG 300
DB 3397 CTAATCAAGTTTTTTGGGTCGAGGTGCGTAAAGCACTAAATCGGAACCTTAAAGGAG 3338
QY 301 CCCCCGATTTAGAGCTTTGACCGGGGAAAGCCGGGGAACGTTGGCGAGAAAGGAAGAA 360
DB 3337 CCCCCGATTTAGAGCTTTGACCGGGGAAAGCCGGGGAACGTTGGCGAGAAAGGAAGAA 3278
QY 361 ACGGAAGAGCGCGGCTAGGCGCTAGGCAAGTGTAGCGGTACGCTCGCGTAAACAC 420
DB 3277 ACGGAAGAGCGCGGCTAGGCGCTAGGCAAGTGTAGCGGTACGCTCGCGTAAACAC 3218

QY 421 CACACCCGCGGCTTAATAGCGCCGCTACAGGGCGGTCCTCCATTCGCCATTTCAGGCTCGC 480
DB |||||
3217 CACACCCGCGGCTTAATAGCGCCGCTACAGGGCGGTCCTCCATTCGCCATTTCAGGCTCGC 3158
QY 481 CAACTGTTGGGAAGGCGGATCGGTGCGGCGCTCTTCCTGCTATTAACGCCAGCTGGCGAAAGG 540
DB |||||
3157 CAACTGTTGGGAAGGCGGATCGGTGCGGCGCTCTTCCTGCTATTAACGCCAGCTGGCGAAAGG 3098
QY 541 GGGATGCTGCTCAAGCGCATTAAGTTGGGTAAAGCCAGGCTTTTCCAGTACAGGCTTG 600
DB |||||
3097 GGGATGCTGCTCAAGCGCATTAAGTTGGGTAAAGCCAGGCTTTTCCAGTACAGGCTTG 3038
QY 601 TAAACAGCGGCGGCTGAGCGGCTCGCTTCACTTCACTGTTTGAACCCGTCGGAGCGG 660
DB |||||
3037 TAAACAGCGGCGGCTGAGCGGCTTAACGACTCACTATAGGGCGAATTGGAGCTCCA 2978
QY 661 GCAGACTCGGGTCAAAATGTTTTTACAGGCTGATGAGCAGATGAAGATGCTCGACAC 720
DB |||||
2977 CCGCGGTGGCGGCGCTCTAGAACTAGTGGATCCCGCGGCTGCAGGAATTCTTATTTGT 2918
QY 721 GCTGCAGAACCGCAGCTAGATTAACTTAACTTAAAGATATATATGAGCTAGCTTAA 780
DB |||||
2917 ATAGTTTCAATGCAATGCTATGTTAATCCAGAGCAGTTTACAACTCAAGAAAGGACCATGT 2858
QY 781 AGATAATC-ATCGCTGTAATTAACCTTAACTGTAATACGACTCACTATAGGGCGAA 839
DB |||||
2857 GGTACGCTTTTCGTTGGGATCTTTCGAAAGGCGAGATTGTCGACAGGTAATGGTTGT 2798
QY 840 TTGGGTACCGGGCCCCCTCGAGTTCGACGGTATCGATAGCTTGAATCGAAATTCCTG 899
DB |||||
2797 CTGGTAAAGGACAGGGCCATCGCCAAATGGAGTATTTGTTGATAATGGTCTGCTAGTT 2738
QY 900 CAGCCCGGGGATCCACTAGTTCTAGAGGGCGGCCCGGGTGGAGCTCCAGCTTTTG 959
DB |||||
2737 GAACGGATCCATCTTCAATGTTGGCGAAATTTGAAAGTTAGCTTTGATTCATTCCTTT 2678
QY 960 TTCCCTTTAGTGAGGGTTAAATTAGATCCCATCGGTCAAATTTTACGACAGCTATCTTTCTA 1019
DB |||||
2677 GTTTGCTCGCGTGATGATATACATTTGTGTGATTAAGTTGTACTCGAGTTGTGTCGA 2618
QY 1020 GGGTTAATCTAGCTGATCAGATCATATCGTCGGGTCTTTTTTCGGGCTCAGTCAATCGC 1079
DB |||||
2617 GAAATGTTTCCATCTCTTTTAAATCAATACCTTTTAACTCGATAC- ----GATTAAACAAG 2563
QY 1080 CCAAGCTGGCTATCTCGGATCGGGAGGAAGACCGCTGCTTTTCCCGGAGGTT 1139
DB |||||
2562 GGTATCACCTTCAAACTTGACTTCAGCAGCGCTTTGTAGTTCGGTCACTCTTTGAAAGA 2503
QY 1140 GAAGCGGATGGAAAGAGTTTCCGAGGATGACTGCTGCTGCAATGAGCTTTGAGCGAAAA 1199
DB |||||
2502 TATAGTGGTTCTCTGTACATAAACCCTTCGGGATGSCACTTTGAAAAAGTCATGCGGTTT 2443
QY 1200 CGCAGTTTACATGATGATTCGGGAAGGTGCGCATGACGCTTTTAAAGGTTGAACG 1259
DB |||||
2442 CATATGATCCGGATAACAGAAAAAGCATTTGAACACCATGAGAGAAAGTAGTGACAAAGTGT 2383
QY 1260 TTCGTTCAAGGCACCTGGGATACAGTTTCGTGCGGCTTTTCCGACACAGTTCGGATG 1319
DB |||||
2382 TGGCCATGGAACAGGTAGTTTTCAGATAGTGAATAAATTAAGTTGAAGCTTTTCGTA 2323
QY 1320 GTCAGCCGGAAGCGCATCAGCAACCCGGAACAATACCGGCGACAGCCGGAACCTGCGTGCC 1379
DB |||||
2322 TGAGCATCACCTTCACCTCTCCATGACAGAAAAATTTGTGCCATTTAAATCACCATC 2263
QY 1380 GGTGTCAGATTAATGACAGCGGTGCGGCTGGGATATTCGTCAGGAGACGGGTAT 1439
DB |||||
2262 TAAATCAACAAGAT- -----TGGGACAACTCCAGTGAAGATTC 2224
QY 1440 CCTGCTGGATGCCGCAAAATGGACATGGATACCCGCTGAGTTACCCGGCGGCGCT 1499
DB |||||
2223 TTCTCCTTTACTCATCGTATCCAGCTTTTGTTCCTTTAGTGAGGGTTAATTGCGCGCT 2164

QY 1500 TGGCGTAATCATGTGCTATAGCTGTTTCTGTGTGAAATTGTTATCCGCTCACAATTCAC 1559
DB |||||
2163 TGGCGTAATCATGTGCTATAGCTGTTTCTGTGTGAAATTGTTATCCGCTCACAATTCAC 2104
QY 1560 ACAACATAGAGCGGGAAGCATAAAGTGTAAAGCCTGGGTGCCCTAAATGAGTGAGCTAAC 1619
DB |||||
2103 ACAACATAGAGCGGGAAGCATAAAGTGTAAAGCCTGGGTGCCCTAAATGAGTGAGCTAAC 2044
QY 1620 TCACATTAATTTGCGTTGCGCTCACTGCCCGCTTTTCCAGTCGGGAAACCTGTCTGCCAGC 1679
DB |||||
2043 TCACATTAATTTGCGTTGCGCTCACTGCCCGCTTTTCCAGTCGGGAAACCTGTCTGCCAGC 1984
QY 1680 TGCATTAATGAATCGGCCAACCGCGGGGAGAGCGGTTTCCGCTAATGCGGCTCTTCCG 1739
DB |||||
1983 TGCATTAATGAATCGGCCAACCGCGGGGAGAGCGGTTTCCGCTAATGCGGCTCTTCCG 1924
QY 1740 CTTCTCGCTCACTGACTCGCTCGCTCGGTCGCTCGGTCGGCTGGGCGAGGGTATCAGCTC 1799
DB |||||
1923 CTTCTCGCTCACTGACTCGCTCGCTCGGTCGCTCGGTCGGCTGGGCGAGCGGTATCAGCTC 1864
QY 1800 ACTCAAGGCGGTAATACCGGTTATCCACAGAAATCAGGGGATAACCGCAGGAAAGACATGT 1859
DB |||||
1863 ACTCAAGGCGGTAATACCGGTTATCCACAGAAATCAGGGGATAACCGCAGGAAAGACATGT 1804
QY 1860 GAGCAAAAGGCCAGCAAAAGGCCAGGAAACCGTAAAGGCCGCGTTGTGCGGTTTTTCC 1919
DB |||||
1803 GAGCAAAAGGCCAGCAAAAGGCCAGGAAACCGTAAAGGCCGCGTTGTGCGGTTTTTCC 1744
QY 1920 ATAGGCTCCGCGCCCTGACGAGCATCAAAATAACGACGCTCAGTCAAGGTGGCGAA 1979
DB |||||
1743 ATAGGCTCCGCGCCCTGACGAGCATCAAAATAACGACGCTCAGTCAAGGTGGCGAA 1684
QY 1980 ACCCGACAGGACTATAAAGATACAGGCGGTTTCCCGCTGGAAGCTCCCTCGTGCGCTCTC 2039
DB |||||
1683 ACCCGACAGGACTATAAAGATACAGGCGGTTTCCCGCTGGAAGCTCCCTCGTGCGCTCTC 1624
QY 2040 CTGTTCCGACCTCGCGCTTACCAGATACCTGTGCGCTTTCTCCCTTCGGGAAAGCGTG 2099
DB |||||
1623 CTGTTCCGACCTCGCGCTTACCAGATACCTGTGCGCTTTCTCCCTTCGGGAAAGCGTG 1564
QY 2100 GCGTTTCTCATAGCTCAGCTGATGATCTCAGTTGCTGAGTCTGCTCGCTCCAGC 2159
DB |||||
1563 GCGTTTCTCATAGCTCAGCTGATGATCTCAGTTGCTGAGTCTGCTCGCTCCAGC 1504
QY 2160 TGGGCTGTGTGCAAGAACCCCGCTTCAGCCCGACCGCTGCGGCTTTATCCGCTAACTATC 2219
DB |||||
1503 TGGGCTGTGTGCAAGAACCCCGCTTCAGCCCGACCGCTGCGGCTTTATCCGCTAACTATC 1444
QY 2220 GTCCTGAGTCCAAACCCGCTAAGACAGCATTTATCGCATGCGCAGCAGCAGCTGGTAACA 2279
DB |||||
1443 GTCCTGAGTCCAAACCCGCTAAGACAGCATTTATCGCATGCGCAGCAGCAGCTGGTAACA 1384
QY 2280 GGAATAGCAGACGAGGTATGTCAGGCTGCTACAGAGTCTTGAAGTGGTGGCTTAAC 2339
DB |||||
1383 GGAATAGCAGACGAGGTATGTCAGGCTGCTACAGAGTCTTGAAGTGGTGGCTTAAC 1324
QY 2340 ACGGCTACACTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCAGCTTACTTCG 2399
DB |||||
1323 ACGGCTACACTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCAGCTTACTTCG 1264
QY 2400 GAAAAAGAGTTGGTAGCTCTTTGATCCCGCAACAAACCAACCGCTGGTGGTGGTTTTT 2459
DB |||||
1263 GAAAAAGAGTTGGTAGCTCTTTGATCCCGCAACAAACCAACCGCTGGTGGTGGTTTTT 1204
QY 2460 TTGTTTCCAGCAGCATGATTAAGCGCAGAAAAAGGATCTCAAGAGATCTCTTTGATCT 2519
DB |||||
1203 TTGTTTGAAGCAGCATGATTAAGCGCAGAAAAAGGATCTCAAGAGATCTCTTTGATCT 1144
QY 2520 TTTTCTAGGGGCTGAGCTCAGTGGAAACGAAAACTCAGTTTAAAGGGATTTTGGTTCATGA 2579
DB |||||
1143 TTTTCTAGGGGCTGAGCTCAGTGGAAACGAAAACTCAGTTTAAAGGGATTTTGGTTCATGA 1084
QY 2580 GATTATCAAAAAAGGATCTTACCTAGATCTCTTTTAAATTAATAAGTAAATCA 2639

Db 1083 GATTATCAAAAGGATCTTCACTAGATCCTTTTAAATTAATAAGTCTTTAAATCAA 1024
QY 2640 TCTAAAGTATATAGTAAATCTTGGTCTGACAGTTTACCAATGCTTAATCAGTGAGGCAC 2699
Db 1023 TCTAAAGTATATAGTAAATCTTGGTCTGACAGTTTACCAATGCTTAATCAGTGAGGCAC 964
QY 2700 CTATCTCAGCGATCTGTCTTATTTGGTTCTCATAGTTGCTGCTGATCTCCCGCTGCTGTAGA 2759
Db 963 CTATCTCAGCGATCTGTCTTATTTGGTTCTCATAGTTGCTGCTGATCTCCCGCTGCTGTAGA 904
QY 2760 TAACTACGATACGGGAGGCTTACCATCTGCGCCCAAGTCTGCAATGATACCGCAGAC 2819
Db 903 TAACTACGATACGGGAGGCTTACCATCTGCGCCCAAGTCTGCAATGATACCGCAGAC 844
QY 2820 CACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCAGCAGCGGAGGCGGAGCGCA 2879
Db 843 CACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCAGCAGCGGAGGCGGAGCGCA 784
QY 2880 GAAGTGGTCTCGCACTTTATCCGCTCATCCAGTCTAATTAATTTGTCGGGGAAGCTA 2939
Db 783 GAAGTGGTCTCGCACTTTATCCGCTCATCCAGTCTAATTAATTTGTCGGGGAAGCTA 724
QY 2940 GAGTAAGTGTGCGCAGTTATAGTTTGGCGCAAGTGTGTCCTGCTTGCATGCTACAGGCATCG 2999
Db 723 GAGTAAGTGTGCGCAGTTATAGTTTGGCGCAAGTGTGTCCTGCTTGCATGCTACAGGCATCG 664
QY 3000 TGGTGTACGCTCGTCTGTTGGTATGGCTTCAATTCAGCTCCGCTCCCAACCATCAAGC 3059
Db 663 TGGTGTACGCTCGTCTGTTGGTATGGCTTCAATTCAGCTCCGCTCCCAACCATCAAGC 604
QY 3060 GAGTTACATGATCCCCCATTTGTGTGCAAAAAAGCGGTAGCTCTTTCGGTCTCCGATCG 3119
Db 603 GAGTTACATGATCCCCCATTTGTGTGCAAAAAAGCGGTAGCTCTTTCGGTCTCCGATCG 544
QY 3120 TTGTGAGAATGAGTTGGCGCAGTGTATCACTCATGTTTATGGCAGCAGCTGCATAATT 3179
Db 543 TTGTGAGAATGAGTTGGCGCAGTGTATCACTCATGTTTATGGCAGCAGCTGCATAATT 484
QY 3180 CTCTTACTGTCTATGCCATCCGTAAGATGCTTTTCTGTGATGCTGAGTACTCAACCAAGT 3239
Db 483 CTCTTACTGTCTATGCCATCCGTAAGATGCTTTTCTGTGATGCTGAGTACTCAACCAAGT 424
QY 3240 CATTCGAGATGATGTATGCGCGCAGCGAGTTGCTCTTTCGGCGGTCAATACGGGATA 3299
Db 423 CATTCGAGATGATGTATGCGCGCAGCGAGTTGCTCTTTCGGCGGTCAATACGGGATA 364
QY 3300 ATACCGCGCAGATAGCAGAACTTTTAAAGTGTCTCATCTTGGTAAACGTTCTTTCGGGCG 3359
Db 363 ATACCGCGCAGATAGCAGAACTTTTAAAGTGTCTCATCTTGGTAAACGTTCTTTCGGGCG 304
QY 3360 GAAAACTCTCAAGGATCTTTACCGTGTTCAGATCCAGTTCGATGTAAACCACTCGTGCAC 3419
Db 303 GAAAACTCTCAAGGATCTTTACCGTGTTCAGATCCAGTTCGATGTAAACCACTCGTGCAC 244
QY 3420 CCAACTGATCTTTCAGCATCTTTTACCTTTTACCGAGCGTTTTCGGGTGAGCAAAAACAGGAA 3479
Db 243 CCAACTGATCTTTCAGCATCTTTTACCTTTTACCGAGCGTTTTCGGGTGAGCAAAAACAGGAA 184
QY 3480 GGCAAAATCCCGCAAAAAGGAATTAAGGCGACACGGAATTTGTAATCTCATCTCT 3539
Db 183 GGCAAAATCCCGCAAAAAGGAATTAAGGCGACACGGAATTTGTAATCTCATCTCT 124
QY 3540 TCCTTTTTCATATTTTCAAGCATTTATCAGGTTTATGCTCTCATGAGCGGATACATAT 3599
Db 123 TCCTTTTTCATATTTTCAAGCATTTATCAGGTTTATGCTCTCATGAGCGGATACATAT 64
QY 3600 TTGAATGTATTTAGAAAAATAAACAATAGGGGTTCCCGGCACATTTTCCCGGAAAGTGC 3659
Db 63 TTGAATGTATTTAGAAAAATAAACAATAGGGGTTCCCGGCACATTTTCCCGGAAAGTGC 4
QY 3660 CAC 3662
|||

Db 3 CAC 1
RESULT 4
US-10-280-913A-32/c
; Sequence 32, Application US/10280913A
; Publication No. US20040110130A1
; GENERAL INFORMATION:
; APPLICANT: Large Scale Biology Corporation
; TITLE OF INVENTION: POPULATION OF POLYNUCLEOTIDE SEQUENCE VARIANTS
; FILE REFERENCE: LSBC-PIG5392-CIP
; CURRENT APPLICATION NUMBER: US/10/280,913A
; CURRENT FILING DATE: 2002-10-25
; PRIOR APPLICATION NUMBER: 60/402,342
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 10/066,390
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 60/268,785
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/266,386
; PRIOR FILING DATE: 2002-02-02
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32
; LENGTH: 3637
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Encodes Cycle 3 BFP gene
US-10-280-913A-32

Query Match 72.8%; Score 2665.4; DB 18; Length 3637;
Best Local Similarity 83.7%; Pred. No. 0;
Matches 3065; Conservative 0; Mismatches 571; Indels 27; Gaps 3;
QY 1 CTAATTCGTAACGGTTAATATTTTGTAAAAATTCGGTTAAATTTTGTAAATCAAGCTC 60
Db 3637 CTAATTCGTAACGGTTAATATTTTGTAAAAATTCGGTTAAATTTTGTAAATCAAGCTC 3578
QY 61 ATTTTAAACCAATAGGCGGAAATCGGCAAAATCCCTTAATAATCAAAAGNATAGACCGA 120
Db 3577 ATTTTAAACCAATAGGCGGAAATCGGCAAAATCCCTTAATAATCAAAAGNATAGACCGA 3518
QY 121 GATAGGTTGAGTGTGTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGAAACGTGACTC 180
Db 3517 GATAGGTTGAGTGTGTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGAAACGTGACTC 3458
QY 181 CAACTCAAGGCGGAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCAC 240
Db 3457 CAACTCAAGGCGGAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCAC 3398
QY 241 CTAATCAAGTTTGTGGGTTCGAGGTTCGTAAGACACTAAATCGGAACCTTAAAGGGAG 300
Db 3397 CTAATCAAGTTTGTGGGTTCGAGGTTCGTAAGACACTAAATCGGAACCTTAAAGGGAG 3338
QY 301 CCCCCGATTTAGAGCTTGACGGGGAAGCCGCGAACAGTGGCGAGAAAGGAGGAAGAA 360
Db 3337 CCCCCGATTTAGAGCTTGACGGGGAAGCCGCGAACAGTGGCGAGAAAGGAGGAAGAA 3278
QY 361 AGCGAAGAGAGCGGGCTTAGGGCGCTAGCGAGTGTAGCGGTACCGCTGGCGTAAACCAC 420
Db 3277 AGCGAAGAGAGCGGGCTTAGGGCGCTAGCGAGTGTAGCGGTACCGCTGGCGTAAACCAC 3318
QY 421 CACACCGCGCGCTTAATGCGCGCTACAGGGCGGTCCCAATTCGCCATTTCAGGCTGCG 480
Db 3217 CACACCGCGCGCTTAATGCGCGCTACAGGGCGGTCCCAATTCGCCATTTCAGGCTGCG 3158
QY 481 CAACTGTTGGGAAGGCGGATCGGTGCGGCTCTTTGCTATTACGCCAGCTGCGCAAGG 540
Db 3157 CAACTGTTGGGAAGGCGGATCGGTGCGGCTCTTTGCTATTACGCCAGCTGCGCAAGG 3098
QY 541 GGGATGTGTCGAAGCGGATTAAGTTGGGTAAAGCCAGGCTTTTCCAGTCAGGAGTTG 600
|||

Dd 3097 GGGATGTGCTCAAGCGCGAATTAAGTTGGGTAAAGCCACAGGGTTTCCACGTCAACGACGTG 3038
Qy 601 TAAACAGCGCCAGTGTAGCGCGCTCGTTCACTTCAACGTTTTTGAACCCGCTGGAGACGG 660
Dd 3037 TAAACAGCGCCAGTGTAGCGCGCTCAATACGACTCACTATAGGCGGAATTGGAGCTCCA 2978
Qy 661 GCAGACTCGCGGTGCAAAATGTGTTTTACAGCGTGTATGGAGCAGATGAAGATGCTCGACAC 720
Dd 2977 CCGCGTGGCGCGCTCTAGACTAGTGGATCCCGCGGCTGCAGGAATTCCTTATTTGT 2918
Qy 721 GCTGCAGAACCGCAGCTAGATTAACCTAGAAAGATAATCATATGTGTGAGTACGTTAA 780
Dd 2917 ATAGTTTCAATCCATGTCATGTGTAATCCCGCAGCAGAGTTTACAACTCAAGAAAGGACCATGT 2858
Qy 781 AGATAATC-ATGCGTAAATTTGACGCAATGGGATCTGTAAATACGACTCACTATAGGCGAA 839
Dd 2857 GGTCAACGCTTTTCGTTGGGATCTTTTGGAAAGGGCAGATTGTGTGCACAGGAATGGTTGT 2798
Qy 840 TTGGGTACCGGCGCCCGCTCGAGGTGCAGCGTATCGATAAGCTTTGATATCGAATTTCCGTG 899
Dd 2797 CTGGTAAAGGACAGGGCCATCGCCAAATTGGAGTATTTTGTGATAATGGTCTGCTAGTT 2738
Qy 900 CAGCCCGGGGATCCACTAGTTTCTAGAGCGCGCCACCGCGGTGGAGCTCCAGCTTTTGTG 959
Dd 2737 GAAAGGATCCATCTTCAATGTGTGGCGAAATTTTGAAGTTAGCTTTGATTCATTTCTTTT 2678
Qy 960 TTCCCTTTAGTGAGGTTAAATTAGATCCATCGCTCAATTTTACGCACTATCTTTCTA 1019
Dd 2677 GTTTGTCTGCGGTGATGATATCATATTGTGTGATTTAAAGTTGTACTCGAGTTTGTGTCGA 2618
Qy 1020 GGGTTAACTAGCTCATCAGGATCATATCGTCGGGTCTTTTTCCGGCTCAGTCATCGC 1079
Dd 2617 GAAATGTTTCCATCTCTCTTTAAATCAATACCTTTTAACTCGATAC-----GATTAACAAG 2563
Qy 1080 CCAAGCTGGGTATCTCGGCATTCGGGAGTGGGAGGAAGACCGCTGTCCTTTTCCCGGAGGTT 1139
Dd 2562 GGTATCACTTCAAACTTTGACTTTCAGCACGGCTCTGTAGTTCCCGTCACTTTTGAAGA 2503
Qy 1140 GAAAGCGCATGGAAGAGTTTGGCGAGGATGACTGCTGCTGATTTGACGTTGAGCGAATA 1199
Dd 2502 TATAGTGGCTTCTGTGATCATTAACCTTTCGGGATGGCACTCTTGAATAAAGTCATGCCGTTT 2443
Qy 1200 CGCAGTTTACCATGATGATTTCCGGAAGGTGTGGCCATTCGACGCTTTTAAACGGTGAAC 1259
Dd 2442 CATATGATCCGGATAACAGAAAGCAATTGAACACCATGAGAGAAAGTAGTGACAAGTGT 2383
Qy 1260 TTGTTTACGGCCACTGCTGGATACAGTTTCTGTCGGGCTTTTCCCGGACACAGTTCCGGATG 1319
Dd 2382 TGGCCATGGAACAGGTAGTTTTCAGTAGTGTGCAATAAAATTTAAGTGAAGCTTTTCCGTA 2323
Qy 1320 GTCAGCCGAGCGCATCAGCAACCGGAACAATACCGGCGACAGCCGGAACCTGCGTGCC 1379
Dd 2322 TGTAGCATCACTTCAACCTCTCCACTGACAGAAATTTGTGCCCATTAACATCAACATC 2263
Qy 1380 GGTGTGAGATTAATGACAGCGGTGCGCGCTGGGATATTACGTCAAGCGAGGACGGGTAT 1439
Dd 2262 TAAATCAACAGAT-----TGGGACAACTCCAGTGAAAAGTTT 2224
Qy 1440 CCTGGCTGGATGCCCGCAGAAATGGAATGGAATPACCCCGTAGTTTACCCGCGGGCGCGCT 1499
Dd 2223 TTCTCTTTACTCATCGGTACCCAGCTTTTGTTCCTTTTAGTGGGGTTTAATTGCGCGCT 2164
Qy 1500 TGGGTTATCATGTTCACTAGCTGTTTCTGTGTGAAATTTGATTCGCTCAATTTCCAC 1559
Dd 2163 TGGCGTAAATCATGTTCACTAGCTGTTTCTGTGTGAAATTTGATTCGCTCAATTTCCAC 2104
Qy 1560 ACAACATACGAGCCGGAAGCATAAAGTGTAAAGCTTGGGGTGCCTTAATGAGTGAGCTTAAC 1619
Dd 2103 ACAACATACGAGCCGGAAGCATAAAGTGTAAAGCTTGGGGTGCCTTAATGAGTGAGCTTAAC 2044
Qy 1620 TCACATTAATTTGCGTTGCGCTCACTGCGCGCTTTTCCAGTCGGGAAACCTGTGTCGCCAGC 1679
Dd 2043 TCACATTAATTTGCGTTGCGCTCACTGCGCGCTTTTCCAGTCGGGAAACCTGTGTCGCCAGC 1984

Qy 1680 TGCATTAATGAATCGGCCAACCGCGGGAGAGCGGTTTTCGTATTGGGCGCTCTTCGG 1739
Dd 1983 TGCATTAATGAATCGGCCAACCGCGGGAGAGCGGTTTTCGTATTGGGCGCTCTTCGG 1924
Qy 1740 CTTTCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 1799
Dd 1923 CTTTCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 1864
Qy 1800 ACTCAAGGCGGTAAATAGGTTATCCACAGAAATCAGGGGATAACCGCAGGAAGACATGT 1859
Dd 1863 ACTCAAGGCGGTAAATAGGTTATCCACAGAAATCAGGGGATAACCGCAGGAAGACATGT 1804
Qy 1860 GAGCAAAAGGCGCAGCAAAAGGCCAGGAACCGTAAAGGCGCGTTCGTGCGGTTTTTCC 1919
Dd 1803 GAGCAAAAGGCGCAGCAAAAGGCCAGGAACCGTAAAGGCGCGTTCGTGCGGTTTTTCC 1744
Qy 1920 ATAGGCTCCGCCCCCTGACGAGCATCAAAATAATCGACGTCAAGTCAGAGGTGGCGAA 1979
Dd 1743 ATAGGCTCCGCCCCCTGACGAGCATCAAAATAATCGACGTCAAGTCAGAGGTGGCGAA 1684
Qy 1980 ACCCGACAGGACTATAAGATACCAAGCGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTC 2039
Dd 1683 ACCCGACAGGACTATAAGATACCAAGCGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTC 1624
Qy 2040 CTGTTCCGACCCCTGCGCTTACCGGATACCTGTCCGCTTCTCCCTTCGGGAAGCGTGG 2099
Dd 1623 CTGTTCCGACCCCTGCGGTTACCGGATACCTGTCCGCTTCTCCCTTCGGGAAGCGTGG 1564
Qy 2100 CGCTTTCTCATPAGCTACGCTGTAGGTATCTCAGTTCCGTTGAGTTCGTTCCGCTCAAGC 2159
Dd 1563 CGCTTTCTCATPAGCTACGCTGTAGGTATCTCAGTTCCGTTGAGTTCGTTCCGCTCAAGC 1504
Qy 2160 TGGGCTGTGTGCAAGAACCCCGCTTCAGCCGAGCCGCTCGCCTTATCCGCTACTATC 2219
Dd 1503 TGGGCTGTGTGCAAGAACCCCGCTTCAGCCGAGCCGCTCGCCTTATCCGCTACTATC 1444
Qy 2220 GTCTTGTAGTCCAAACCGGTAAAGACACGACTTATCGCCACTGCGCAGCAGCCACTGCTGAACA 2279
Dd 1443 GTCTTGTAGTCCAAACCGGTAAAGACACGACTTATCGCCACTGCGCAGCAGCCACTGCTGAACA 1384
Qy 2280 GGATTTAGCAGACGAGGTATGTAGCGGTGTCTACAGAGTTCTTTGAAGTGGTGGCCTAAT 2339
Dd 1383 GGATTTAGCAGACGAGGTATGTAGCGGTGTCTACAGAGTTCTTTGAAGTGGTGGCCTAAT 1324
Qy 2340 ACGGCTACACTAGAGGACGACTATTTGGTATCTCGCTCTGCTGAAGCAGTTACTTCTCG 2399
Dd 1323 ACGGCTACACTAGAGGACGACTATTTGGTATCTCGCTCTGCTGAAGCAGTTACTTCTCG 1264
Qy 2400 GAAAAAGAGTTGGTAGCTTCTTGATCCGGCAAAACAAACACCGCTGCTGAGCGTGGTTTTT 2459
Dd 1263 GAAAAAGAGTTGGTAGCTTCTTGATCCGGCAAAACAAACACCGCTGCTGAGCGTGGTTTTT 1204
Qy 2460 TTGTTTGCAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAGATCCTTTTGATCT 2519
Dd 1203 TTGTTTGCAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAGATCCTTTTGATCT 1144
Qy 2520 TTTTCTACGGGCTGTGACGCTCAGTTGGAGCAAAACTCAGTTTAAAGGATTTTGGTCACTGA 2579
Dd 1143 TTTTCTACGGGCTGTGACGCTCAGTTGGAGCAAAACTCAGTTTAAAGGATTTTGGTCACTGA 1084
Qy 2580 GATTATCAAAAAGGATCTTCACTAGATCTTTTAAATTAATAAATAAAGTTTTAAATCAA 2639
Dd 1083 GATTATCAAAAAGGATCTTCACTAGATCTTTTAAATTAATAAATAAAGTTTTAAATCAA 1024
Qy 2640 TCTAAAGTATATATAGTAAACTTTGGTCTGACAGTTACAAATGCTTAATCATGTAGGAGC 2699
Dd 1023 TCTAAAGTATATATAGTAAACTTTGGTCTGACAGTTACAAATGCTTAATCATGTAGGAGC 964
Qy 2700 CTATCTCAGCGCATCTGCTATTTTCGTTTCATCCATAGTTGCTGACTCCCGCTCGGTGAGA 2759
Dd 963 CTATCTCAGCGCATCTGCTATTTTCGTTTCATCCATAGTTGCTGACTCCCGCTCGGTGAGA 904


```
QY 2760 TAACTACGATCGGAGGGCTTACCATCTGGCCCCAGTCTGCAATGATACCGCGAGACC 2819
Db 903 TAACTACGATCGGAGGGCTTACCATCTGGCCCCAGTCTGCAATGATACCGCGAGACC 844
QY 2820 CACGCTCACGGCTCCAGATTTATCAGCAATAAACCCAGCAGCCGGAAGGGCGGAGCGCA 2879
Db 843 CACGCTCACGGCTCCAGATTTATCAGCAATAAACCCAGCAGCCGGAAGGGCGGAGCGCA 784
QY 2880 GAAGTGGTCTGCAACTTTATCCGCTCCATCCAGTCTTAATTAATTTGTCGGGAGCTTA 2939
Db 783 GAAGTGGTCTGCAACTTTATCCGCTCCATCCAGTCTTAATTAATTTGTCGGGAGCTTA 724
QY 2940 GAGTAAGTAGTTCCGCGAGTTAATTTGCGGCAAGTTGTCGCAATGCTACAGGATCG 2999
Db 723 GAGTAAGTAGTTCCGCGAGTTAATTTGCGGCAAGTTGTCGCAATGCTACAGGATCG 664
QY 3000 TGGTGTCCAGCTCGCTGCTGTTGATGCTTTCATTCAGCTCCGTTCCCAACCATCAAGC 3059
Db 663 TGGTGTCCAGCTCGCTGCTGTTGATGCTTTCATTCAGCTCCGTTCCCAACCATCAAGC 604
QY 3060 GAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGTTAGCTCTCTTCGGTCTCCGATCG 3119
Db 603 GAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGTTAGCTCTCTTCGGTCTCCGATCG 544
QY 3120 TTGTCGAGAAAGTAGTTGCGCGAGTGTATCACTCATGTTATGCGAGCACTGCATAATT 3179
Db 543 TTGTCGAGAAAGTAGTTGCGCGAGTGTATCACTCATGTTATGCGAGCACTGCATAATT 484
QY 3180 CTCCTACTGTCATGCCATCCGTAAGATGCTTTCGTGATCTGTGATGACTCAACCAAGT 3239
Db 483 CTCCTACTGTCATGCCATCCGTAAGATGCTTTCGTGATCTGTGATGACTCAACCAAGT 424
QY 3240 CATTTCTGAGAAATGATGTATGCGCGACCGAGTTGCTCTTTCGCGCGGTCAATACGGGATA 3299
Db 423 CATTTCTGAGAAATGATGTATGCGCGACCGAGTTGCTCTTTCGCGCGGTCAATACGGGATA 364
QY 3300 ATACGCGGCCACATAGCAGAACTTTTAAAGTGTCTATCATTTGGAAGAGCTTCTTGGGCG 3359
Db 363 ATACGCGGCCACATAGCAGAACTTTTAAAGTGTCTATCATTTGGAAGAGCTTCTTGGGCG 304
QY 3360 GAAACTCTCAAGGATCTTACCGCTGTTCAGATCCAGTTTCGATGTAACCCACTCTGTGCAC 3419
Db 303 GAAACTCTCAAGGATCTTACCGCTGTTCAGATCCAGTTTCGATGTAACCCACTCTGTGCAC 244
QY 3420 CCAACTGATCTTTCAGCATCTTTTACTTTTCAACAGCGTTTCTGGTGAGCAAAAAACAGGAA 3479
Db 243 CCAACTGATCTTTCAGCATCTTTTACTTTTCAACAGCGTTTCTGGTGAGCAAAAAACAGGAA 184
QY 3480 GGCAAAATGCCGCAAAAAAGGAATAAGGGCGACACAGGAAATGTTGAATATCTACTCTCT 3539
Db 183 GGCAAAATGCCGCAAAAAAGGAATAAGGGCGACACAGGAAATGTTGAATATCTACTCTCT 124
QY 3540 TCCTTTTCAATATTTAAGCATTATACAGGTTATGTTCTCATGAGCGGATACATAT 3599
Db 123 TCCTTTTCAATATTTAAGCATTATACAGGTTATGTTCTCATGAGCGGATACATAT 64
QY 3600 TTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCGGAAAGTGC 3659
Db 63 TTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCGGAAAGTGC 4
QY 3660 CAC 3662
Db 3 CAC 1
```

RESULT 5
US-10-684-134-32/c
; Sequence 32, Application US/10684134
; Publication No. US2004014243A1
; GENERAL INFORMATION:
; APPLICANT: Large Scale Biology Corporation
; TITLE OF INVENTION: POPULATION OF POLYNUCLEOTIDE SEQUENCE VARIANTS
; FILE REFERENCE: LSBC-PLG5392-CIP

```
; CURRENT APPLICATION NUMBER: US/10/684,134  
; CURRENT FILING DATE: 2003-10-10  
; PRIOR APPLICATION NUMBER: 60/402,342  
; PRIOR FILING DATE: 2002-08-08  
; PRIOR APPLICATION NUMBER: 10/066,390  
; PRIOR FILING DATE: 2002-02-01  
; PRIOR APPLICATION NUMBER: 60/268,785  
; PRIOR FILING DATE: 2001-02-14  
; PRIOR APPLICATION NUMBER: 60/266,386  
; PRIOR FILING DATE: 2002-02-02  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 32  
; LENGTH: 3637  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Encodes Cycle 3 BFP gene  
US-10-684-134-32  
  
Query Match 72.8%; Score 2665.4; DB 18; Length 3637;  
Best Local Similarity 83.7%; Pred. No. 0;  
Matches 3065; Conservative 0; Mismatches 571; Indels 27; Gaps 3;  
  
QY 1 CTAATTTGTAACGGTTAATAATTTGTTAAAAATTCGGTTAAAAATTTTGTGTAATCAGCTC 60  
Db 3637 CTAATTTGTAACGGTTAATAATTTGTTAAAAATTCGGTTAAAAATTTTGTGTAATCAGCTC 3578  
  
QY 61 ATTTTAAACCAATAGCCGGAATCGCAAAATCCCTTATAATCAAAAGAAATAGACCGA 120  
Db 3577 ATTTTAAACCAATAGCCGGAATCGCAAAATCCCTTATAATCAAAAGAAATAGACCGA 3518  
  
QY 121 GATAGGTTGAGTGTGTTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGAAACGTGGACTC 180  
Db 3517 GATAGGTTGAGTGTGTTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGAAACGTGGACTC 3458  
  
QY 181 CAACTGTAAGGGCGGAAACCGTCTATCAGGGCGATGCGCCACTACGTGAACCATCACC 240  
Db 3457 CAACTGTAAGGGCGGAAACCGTCTATCAGGGCGATGCGCCACTACGTGAACCATCACC 3398  
  
QY 241 CTAATCAAGTTTTTGGGGTCGAGGTGCGTAAAGCACTAAATCGGAACCCCTAAAGGGAG 300  
Db 3397 CTAATCAAGTTTTTGGGGTCGAGGTGCGTAAAGCACTAAATCGGAACCCCTAAAGGGAG 3338  
  
QY 301 CCCCCGATTTAGAGCTTTGACGGGGAAAGCCGGCGAAGCTGGCGAGAAAGGAGGGAAGAA 360  
Db 3337 CCCCCGATTTAGAGCTTTGACGGGGAAAGCCGGCGAAGCTGGCGAGAAAGGAGGGAAGAA 3278  
  
QY 361 AGCGAAAGAGCGGCGCTAGGGCGTGGCAAGTGTAGCGGTCACTCGCTGGCGTAAACAC 420  
Db 3277 AGCGAAAGAGCGGCGCTAGGGCGTGGCAAGTGTAGCGGTCACTCGCTGGCGTAAACAC 3218  
  
QY 421 CACACCCGCGCTTAATGCGCCCTACAGGGCGGTCCCATTCGCCATTTCAGGCTGCG 480  
Db 3217 CACACCCGCGCTTAATGCGCCCTACAGGGCGGTCCCATTCGCCATTTCAGGCTGCG 3158  
  
QY 481 CAACTGTTGGGAAGGGCGATCCGTCGGGCTCTTCGCTATTACGCCAGTGGCGAAAGG 540  
Db 3157 CAACTGTTGGGAAGGGCGATCCGTCGGGCTCTTCGCTATTACGCCAGTGGCGAAAGG 3098  
  
QY 541 GGGATGTGTCGAAGCGGATTAAGTTGGGTAAACCGCAGGGTTTTTCCAGTCAGAGCTTG 600  
Db 3097 GGGATGTGTCGAAGCGGATTAAGTTGGGTAAACCGCAGGGTTTTTCCAGTCAGAGCTTG 3038  
  
QY 601 TAAACGAGCGCGCTAGCGCGCTTCGTTTACAGCTGATGAGAGTGTGAGAGTGTGAGCTCG 660  
Db 3037 TAAACGAGCGCGCTAGCGCGCTTCGTTTACAGCTGATGAGAGTGTGAGAGTGTGAGCTCG 2978  
  
QY 661 GCAGACTCGCGTCAAAATGTTTATACAGCTGATGAGAGTGTGAGAGTGTGAGCTCG 720  
Db 2977 CGCGGTGCGCGCTCGCTCTAGAACTAGTGTGATCCCGGGCTGCAGGAATTTCTATTGT 2918  
  
QY 721 GCTGCAGAACACGCGAGCTAGATTAAACCCCTAGAAAGATAATCATATTGTGACGTACGTTAA 780
```

Db 2917 ATAGTTCATCCATGTCATGTAATCCAGCAGCAGTACAAACTCAAGAAAGCACCATGT 2858
Qy 781 AGATAATC-ATCGGTAAATAATGACCATGGGATCTGTAATACGACTCACTATAGGGCGAA 839
Db 2857 GGTCAAGCTTTTCGTTGGGATCTTTTCGAAAGGGCAGATGTTGTCACAGAGTAATGGTTGT 2798
Qy 840 TTGGGTACCGGGCCCCCTCGAGGTGACAGGTATCGATAGCTTTGATATCGAAATTCCTG 899
Db 2797 CTGGTAAAGGACAGGGCCATCGCCAAATGGAGTATTTTGTGATAAGTGTCTGTAGTT 2738
Qy 900 CAGCCCGGGGATCACTAGTTCTAGAGCGGCCCAACCGGGTGGAGCTCCAGCTTTTG 959
Db 2737 GAAAGGATCACTCTCAATGTTGTGGCGAAATTTTGAAGTTAGCTTTTGATTCATCTTTT 2678
Qy 960 TTCCCTTTAGTGAGGGTTAAATAGATCCATGCGTCAATTTTACGCAGACTATCTTTCTA 1019
Db 2677 GTTGTCTGCGGTGATATACATTTGTGTGATTTAAAGTTGTACTCGAGTTTGTGTCGA 2618
Qy 1020 GGGTTAACTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGCTCAGTCAATCGC 1079
Db 2617 GAAATGTTCCATCTCTTTTAAATCAATACCTTTTAACTCGATAC-----GATTAAACAAG 2563
Qy 1080 CCAAGCTGGCGTATCTGGGATCGGGAGGAAGACCGGCTTTTTCGGCGAGGTT 1139
Db 2562 GGTATCACTTCAAACTTTGACTTCAGCACGGCTCTGTAGTTCCCGTCATCTTTGAAAGA 2503
Qy 1140 GAAGCGGATGGAAGAGTTTCCCGAGGATGACTGCTGCTGCATTTGAGTTCAGCGAAAA 1199
Db 2502 TATAGTGGTTCCCTGTACATAACCTTCGGGCAATGGCACTTTGAAAAAGTCAATGCCGTTT 2443
Qy 1200 CGCAGTTTACCATGATGATTCGGGAAGTGTGGCCATGACAGCCCTTTTAAAGCGTGAAC 1259
Db 2442 CATATGATCCGGATAACGAGAAAGCATTTGAACCATGAGAGAAAGTAGTGACAAAGTGT 2383
Qy 1260 TTGCTTACGGCCACTGGGATACCAAGTTCTGTGGGGCTTTTTCGGACACAGTTCCGGATG 1319
Db 2382 TGGCCATGGAACAGGTAGTTTTCAGTAGTAGTCAATAAATAATTAAGTGAAGCTTTCCGTA 2323
Qy 1320 GTCAGCCGAGCGCATCAGCAACCGGAACAATACCGGGACAGCGGAACCTGCGTGCC 1379
Db 2322 TGTAGCATCACCTTCAACCTCTCCACTGACAGAAATTTGTGCCATTTAAACATCACCATC 2263
Qy 1380 GGTGTGCAGATTAATGACAGCGGTGCGCGCTGGGATATACGTGAGCGAGGACGGGTAT 1439
Db 2262 TAAATCAACAAGAT-----TGGGACAACTCCAGTGAAAGTTC 2224
Qy 1440 CTTGCTGGATGCCGCAAGAAATGGACATGGATACCCCGTAGTTTACCCGGCGGGCGCT 1499
Db 2223 TTTCTCTTTACTCATCGGTACCCAGCTTTTGTTCCTTTTAGTGAGGGTTAATTTGCGCGCT 2164
Qy 1500 TGGCGTAAATCATGGTCATAGCTTTTCTGTGTGAATTTGTTATCCGCTCACAATTCAC 1559
Db 2163 TGGCGTAAATCATGGTCATAGCTTTTCTGTGTGAATTTGTTATCCGCTCACAATTCAC 2104
Qy 1560 ACAACATACGAGCGGGAAGCATAAAGTGTAAAGCTGGGGTGCCTTAATGAGTGAGCTAAC 1619
Db 2103 ACAACATACGAGCGGGAAGCATAAAGTGTAAAGCTGGGGTGCCTTAATGAGTGAGCTAAC 2044
Qy 1620 TCACATTAATTTGCTGTGGCTCATCTGCCGCTTTCCAGTCGGGAAAACCTGTGTCGCCAGC 1679
Db 2043 TCACATTAATTTGCTGTGGCTCATCTGCCGCTTTCCAGTCGGGAAAACCTGTGTCGCCAGC 1984
Qy 1680 TGCAATTAATGAAATCGGCCAAACCGCGGGAGAGCGGTTTCGGTATTTGGGCGCTCTTCGG 1739
Db 1983 TGCAATTAATGAAATCGGCCAAACCGCGGGAGAGCGGTTTTCGGTATTTGGGCGCTCTTCGG 1924
Qy 1740 CTTTCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 1799
Db 1923 CTTTCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 1864
Qy 1800 ACTCAAAGCGGTAATACGGTTATCCACAGAAATACAGGGATAAACGACGAAAGAACATGT 1859
|||||

Db 1863 ACTCAAAGCGGTAATACGGTTATCCACAGAAATCAGGGGATAACGACGAAAGAACATGT 1804
Qy 1860 GAGCAAAAGGCGCAGCAAAAGCCAGGAACCGTTAAAAAAGCCGCGTGTGCTGGCGTTTTTCC 1919
Db 1803 GAGCAAAAGGCGCAGCAAAAGCCAGGAACCGTTAAAAAAGCCGCGTGTGCTGGCGTTTTTCC 1744
Qy 1920 ATAGGCTCCGGCCCCCTTGACGAGCATCAAAAAATCGACGCTCAAGTTCAGAGGTGGCGAA 1979
Db 1743 ATAGGCTCCGGCCCCCTTGACGAGCATCAAAAAATCGACGCTCAAGTTCAGAGGTGGCGAA 1684
Qy 1980 ACCGAGAGGACTATAAAGATACAGAGGCTTTCCCTCGGAAGCTCCCTCGTGCCTCTC 2039
Db 1683 ACCGAGAGGACTATAAAGATACAGAGGCTTTCCCTCGGAAGCTCCCTCGTGCCTCTC 1624
Qy 2040 CTGTTCCGACCTCGCGCTTACCGGATACCTGTCGCGCTTTCTCCCTTCGGGAACGCTGG 2099
Db 1623 CTGTTCCGACCTCGCGCTTACCGGATACCTGTCGCGCTTTCTCCCTTCGGGAACGCTGG 1564
Qy 2100 CGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCCGGTGTAGGTCTGCTCCCAAGC 2159
Db 1563 CGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCCGGTGTAGGTCTGCTCCCAAGC 1504
Qy 2160 TGGGCTGTGCAAGAACCCCGCTTACGCCGACCGCTGGCTTTATCCGTAATCTATC 2219
Db 1503 TGGGCTGTGTCACGAACCCCGCTTACGCCGACCGCTGGCTTTATCCGTAATCTATC 1444
Qy 2220 GTCTTGTAGTCCAAACCGGTAAGACACGACTTATGCCACTGCGACAGCCACTGGTAACA 2279
Db 1443 GTCTTGTAGTCCAAACCGGTAAGACACGACTTATGCCACTGCGACAGCCACTGGTAACA 1384
Qy 2280 GGATTAGCAGACGAGGTATGTAGGCGGTGTACAGAGTTCTTGAAGTGGTGGCCTAACT 2339
Db 1383 GGATTAGCAGACGAGGTATGTAGGCGGTGTACAGAGTTCTTGAAGTGGTGGCCTAACT 1324
Qy 2340 ACGGCTACACTAGAAAGACAGTATTTGGTATCTGGCTCTGCTGAGGCGAGTTACTTCG 2399
Db 1323 ACGGCTACACTAGAAAGACAGTATTTGGTATCTGGCTCTGCTGAGGCGAGTTACTTCG 1264
Qy 2400 GAAAAAGAGTTGTAGTCTTTGATCCGGCAAAACAAACCAACCGCTGCTGAGCGGTGTTTTT 2459
Db 1263 GAAAAAGAGTTGTAGTCTTTGATCCGGCAAAACAAACCAACCGCTGCTGAGCGGTGTTTTT 1204
Qy 2460 TTGTTTGAACGACGAGATTACGCGCAGAAAAAAGGATCTCAAGAGATCTTTTGTATCT 2519
Db 1203 TTGTTTGAACGACGAGATTACGCGCAGAAAAAAGGATCTCAAGAGATCTTTTGTATCT 1144
Qy 2520 TTTCTAGCGGCTGTGAGCTCAGTGGGAACGAAACTCAGCTTAAAGGATTTTGGTCATGA 2579
Db 1143 TTTCTAGCGGCTGTGAGCTCAGTGGGAACGAAACTCAGCTTAAAGGATTTTGGTCATGA 1084
Qy 2580 GATTATCAAAAAAGGATCTTCACTAGATCTTTTAAATTTAAATTTAAATTTAAATCAA 2639
Db 1083 GATTATCAAAAAAGGATCTTCACTAGATCTTTTAAATTTAAATTTAAATTTAAATCAA 1024
Qy 2640 TCTAAAGPATATATAGTAAACTTTGGTCTGACAGTTACCAATGCTTAATTCAGTAGGAC 2699
Db 1023 TCTAAAGPATATATAGTAAACTTTGGTCTGACAGTTACCAATGCTTAATTCAGTAGGAC 964
Qy 2700 CTATCTCAGGATCTGTCTATTTGTTTCATCCATAGTTGCTGACTCCCGCTGTGTA 2759
Db 963 CTATCTCAGGATCTGTCTATTTGTTTCATCCATAGTTGCTGACTCCCGCTGTGTA 904
Qy 2760 TAACTTACGATACGGAGGGCTTTACCTGCGCCCGAGTGTGCAATGATACCGCGAGACC 2819
Db 903 TAACTTACGATACGGAGGGCTTTACCTGCGCCCGAGTGTGCAATGATACCGCGAGACC 844
Qy 2820 CACGCTCACCGGCTCCAGATTTATCAGCAATAAACAGCCAGCCGGAAGGGCCGAGCGCA 2879
Db 843 CACGCTCACCGGCTCCAGATTTATCAGCAATAAACAGCCAGCCGGAAGGGCCGAGCGCA 784
Qy 2880 GAGTGGTCTCGCAACTTTTATCCGCTCCATCCAGTCTTAATTTGTTGGCGGAGCTA 2939
Db 783 GAGTGGTCTCGCAACTTTTATCCGCTCCATCCAGTCTTAATTTGTTGGCGGAGCTA 724
|||||


```
QY 2940 GAGTAAGTAGTTCGCCAGTTAATAGTTTGGCGAACGTTGTTGCCATTGCTTACAGGATCG 2999
Db 723 GAGTAAGTAGTTCGCCAGTTAATAGTTTGGCGAACGTTGTTGCCATTGCTTACAGGATCG 664
QY 3000 TGGTGTCAAGCTCGTCTGTTGGTATGCGTTCATTTCAGCTCCGTTCCCAACGATCAAGGC 3059
Db 663 TGGTGTCAAGCTCGTCTGTTGGTATGCGTTCATTTCAGCTCCGTTCCCAACGATCAAGGC 604
QY 3060 GAGTTACATGATCCCCCATGTTGTGCAAAAAGCGTTAGCTCTTCGTCCTCCGATCG 3119
Db 603 GAGTTACATGATCCCCCATGTTGTGCAAAAAGCGTTAGCTCTTCGTCCTCCGATCG 544
QY 3120 TTGTCAGAAGTAGTTGGCGCAGTGTATACATCATGTTATGCGAGCAGCTGCATAAT 3179
Db 543 TTGTCAGAAGTAGTTGGCGCAGTGTATACATCATGTTATGCGAGCAGCTGCATAAT 484
QY 3180 CTCCTTACTCTCATGCCATCCGTAAGATGCTTTCTGTGACTGGTGAAGTACTCAACCAAGT 3239
Db 483 CTCCTTACTCTCATGCCATCCGTAAGATGCTTTCTGTGACTGGTGAAGTACTCAACCAAGT 424
QY 3240 CATCTTGAGATAGTGTATGCGCGCAGCGAGTTGCTCTTGCCCGCGCTCAATACGGGATA 3299
Db 423 CATCTTGAGATAGTGTATGCGCGCAGCGAGTTGCTCTTGCCCGCGCTCAATACGGGATA 364
QY 3300 ATACCGCGCCACATAGCAGAACTTTTAAAGTGTCTCATCATTTGGAAAACGTTCTTCGGGCG 3359
Db 363 ATACCGCGCCACATAGCAGAACTTTTAAAGTGTCTCATCATTTGGAAAACGTTCTTCGGGCG 304
QY 3360 GAAACTCTCAGGATCTTACCGCTGTGAGATCCAGTTGCTCTTGCCCGCGCTCAATACGGGATA 3419
Db 303 GAAACTCTCAGGATCTTACCGCTGTGAGATCCAGTTGCTCTTGCCCGCGCTCAATACGGGATA 244
QY 3420 CCAACTGATCTTCAGCATCTTTTACTTTTCCACAGCGTTTCTGGTGAGCAAAAACAGGAA 3479
Db 243 CCAACTGATCTTCAGCATCTTTTACTTTTCCACAGCGTTTCTGGTGAGCAAAAACAGGAA 184
QY 3480 GGCATAATCCGCAAAAAGGGAATAAGGGCGACACGGAATAATGTTGAATATCTCATCTCT 3539
Db 183 GGCATAATCCGCAAAAAGGGAATAAGGGCGACACGGAATAATGTTGAATATCTCATCTCT 124
QY 3540 TCCTTTTCAATATTAATGAGCAATTTATCAGGGTTATGTTCTCATGAGGGGATACATAT 3599
Db 123 TCCTTTTCAATATTAATGAGCAATTTATCAGGGTTATGTTCTCATGAGGGGATACATAT 64
QY 3600 TTGAATGTTATTTAGAAAAATAAACAATAAGGGGTTCCGCGCATTTTCCCGAAAAAGTGC 3659
Db 63 TTGAATGTTATTTAGAAAAATAAACAATAAGGGGTTCCGCGCATTTTCCCGAAAAAGTGC 4
QY 3660 CAC 3662
Db 3 CAC 1
```

RESULT 6

```
US-10-637-758-32/c
; Sequence 32, Application US/10637758
; Publication No. US20040180352A1
; GENERAL INFORMATION:
; APPLICANT: Large Scale Biology Corporation
; TITLE OF INVENTION: POPULATION OF POLYNUCLEOTIDE SEQUENCE VARIANTS
; FILE REFERENCE: LSBC-PLG5392-CIP
; CURRENT APPLICATION NUMBER: US/10/637, 758
; CURRENT FILING DATE: 2003-08-08
; PRIOR APPLICATION NUMBER: US/10/280, 913
; PRIOR FILING DATE: 2002-10-25
; PRIOR APPLICATION NUMBER: 60/402, 342
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 10/066, 390
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 60/268, 785
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/266, 386
```

```
; PRIOR FILING DATE: 2002-02-02
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32
; LENGTH: 3637
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Encodes Cycle 3 BFP gene
US-10-637-758-32
```

```
Query Match 72.8%; Score 2665.4; DB 18; Length 3637;
Best Local Similarity 83.7%; Pred. No. 0;
Matches 3065; Conservative 0; Mismatches 571; Indels 27; Gaps 3;

QY 1 CTAAATTGTAAGCGTTAAATATTTTGTAAAAATTCGCGTTAAAAATTTTGTAAATTCAGCTC 60
Db 3637 CTAAATTGTAAGCGTTAAATATTTTGTAAAAATTCGCGTTAAAAATTTTGTAAATTCAGCTC 3578
QY 61 ATTTTAAACCAATAGCGCGAAATCGGCAAAATCCCTTATAAATCAAAAGAAATAGACCGA 120
Db 3577 ATTTTAAACCAATAGCGCGAAATCGGCAAAATCCCTTATAAATCAAAAGAAATAGACCGA 3518
QY 121 GATAGGTTGAGTGTGTTCCAGTTTGGAAACAGAGTCCCACTATTAAAGAACGTGGACTC 180
Db 3517 GATAGGTTGAGTGTGTTCCAGTTTGGAAACAGAGTCCCACTATTAAAGAACGTGGACTC 3458
QY 181 CAACTGCAAAAGGCGGCAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 240
Db 3457 CAACTGCAAAAGGCGGCAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 3398
QY 241 CTAATCAAGTTTTTTGGGTCGAGGTGCGCTAAAGCACTAAATCGGAACCCCTAAAGGGAG 300
Db 3397 CTAATCAAGTTTTTTGGGTCGAGGTGCGCTAAAGCACTAAATCGGAACCCCTAAAGGGAG 3338
QY 301 CCCCAGATTAGAGCTTTGACGGGGAAGCCGGCAACGTTGGCGAGAAAGGAGGAGAA 360
Db 3337 CCCCAGATTAGAGCTTTGACGGGGAAGCCGGCAACGTTGGCGAGAAAGGAGGAGAA 3278
QY 361 AGCGAAAGGAGCGGGCGCTAGGGCGCTGCAAGTGTAGCGGTCACTGCGCGCTAACCCAC 420
Db 3277 AGCGAAAGGAGCGGGCGCTAGGGCGCTGCAAGTGTAGCGGTCACTGCGCGCTAACCCAC 3218
QY 421 CACACCGCGCGCTTAAATGCGCGCTACAGGGCGCTCCCATTCGCAATTCAGGTGCG 480
Db 3217 CACACCGCGCGCTTAAATGCGCGCTACAGGGCGCTCCCATTCGCAATTCAGGTGCG 3158
QY 481 CAACTGTTGGGAAGGGCGATCGGTGCGGGCTCTTCGCTATTACGCCAGCTGGCGAAGG 540
Db 3157 CAACTGTTGGGAAGGGCGATCGGTGCGGGCTCTTCGCTATTACGCCAGCTGGCGAAGG 3098
QY 541 GGGATGTCGCAAGGCGATTAAAGTTGGGTAAAGCCAGAGGTTTCCAGTCACGAGCTTG 600
Db 3097 GGGATGTCGCAAGGCGATTAAAGTTGGGTAAAGCCAGAGGTTTCCAGTCACGAGCTTG 3038
QY 601 TAAACGACGGCGAGTGAAGCGCGCTCTGTTCAATTCATCGTTTTTGAACCCCGTGGAGACGG 660
Db 3037 TAAACGACGGCGAGTGAAGCGCGCTCTGTTCAATTCATCGTTTTTGAACCCCGTGGAGTCCA 2978
QY 661 GCAGACTCGCGTGCAAAATGTTTACAGCGGTGATGGAGCAGATGAAGTGTCTGCAC 720
Db 2977 CCGCGGTGGCGCGCTCTAGAACTAGTGGATCCCGCGGCTGCAGGAATTTCTTATTTGT 2918
QY 721 GCTGCAGAACACGCGAGCTAGATTAAACCTAGAAAAGATAATCATATTTGTGACGTACGTAA 780
Db 2917 ATAGTTTATCCATGCCATGTGTAAATCCAGCAGCAGTTACAAACTCAAGAAAGACCATGT 2858
QY 781 AGATAATC-ATGCGTAAAAATTCAGCGCATGGGATCTGTAATACGACTCACTATATAGGCGAA 839
Db 2857 GGTCAAGCTTTTGGTGGGATCTTTTGAAGGGGAGATTGTTGTCGACAGGTAATGTTGT 2798
QY 840 TTGGGTACCGGGCCCCCTCTCGAGGTGCGAGGTATCGATAAGCTTGTATATCGAATTCCTG 899
Db 840 TTGGGTACCGGGCCCCCTCTCGAGGTGCGAGGTATCGATAAGCTTGTATATCGAATTCCTG 899
```

Db 2797 CTGGTAAAGGACAGGGCCCATCCCAATTGGAGTATTTTGTGATAATGGTCTGCTAGTT 2738
Qy 900 CAGCCGGGGATCCATAGTTCTAGAGCGCGCCACCGCGGTGGAGCTCCAGCTTTTG 959
Db 2737 GAAACGATCCATCTTCAATGTTGTGGCAATTTTGAAGTAGCTTTGATTCATCTCTTT 2678
Qy 960 TTCCCTTTAGTGAGGGTTAAATTTAGATCCCATCGTCAATTTTACGACAGACTATCTTTCTA 1019
Db 2677 GTTGTCTGCCGTGATATACATTGTGTGAGTTAAAGTTGTACTCGAGTTTGTGTCCGA 2618
Qy 1020 GGGTTAACTAGCTGCATCAGGATCATATCGTGGGTCTTTTTCGGGCTCAGTCATCGC 1079
Db 2637 GAATGTTTCCATCTTTCTTTAAATCAATACCTTTTAACTCGATAC-----GATTAAACA 2563
Qy 1080 CCAAGCTGGCGTATCTGGGCATCGGGAGGAAGCCGCTTTTCCCGCAGGTT 1139
Db 2562 GGTATCACCTTCAAACTTGACTTCAGACGCGTCTTTAGTTCCTGTCATCTTTGAAAGA 2503
Qy 1140 GAAGCGCATGGAAGAGTTTSCGAGGATGACTGCTGTCATTTGACGTTTGAGCGAAAA 1199
Db 2502 TATAGTGGTCTCTGTACATAACCTTCGGGCATGGCACTCTTTGAAAAAGTCATGCCGTTT 2443
Qy 1200 CGCAGTTTACCATGATGATTCGGGAAGGTGGGCCATGCAGCCCTTTAAACGTTGAACTG 1259
Db 2442 CATATGATCCGGATAACGAGAAAAAGCATTTGAACACCATGAGAGAAAGTAGTGACAAAGTGT 2383
Qy 1260 TTGCTTACGGCCACCTGGGATACCAAGTTTCGTGCGGCTTTTCCGCACACAGTTCCGGATG 1319
Db 2382 TGGCCATGGAAACAGGTAGTTTTCAGTAGTAGTGCAATAAATTTAAGTGTAAAGCTTCCGTA 2323
Qy 1320 GTACCCGGAAGCGCATCAGCAACCCGAACATACCGGCGACAGCGGAACCTGCCGTGCC 1379
Db 2322 TGTAGCATCACCTTCACCTCTCCACTGCACAGAAAAATTTGTGCCCATTAACATCACCATC 2263
Qy 1380 GGTGTGCAGATTAATGACAGCGGTGCGGCTGGGATATTAGTCAGCGAGGACGGGTAT 1439
Db 2262 TAAATTCACAAAGAAAT-----TGGGCAACTCCAGTGAAGAAAGTTC 2224
Qy 1440 CTTGCTGGATGCGCAGAAATGGACATGGATPACCCCGTGGATTACCCGGCGCGCT 1499
Db 2223 TTCTCCTTTATCTCATCGGTACCCAGCTTTGTTCCTTTTAGTGAGGGTTAATTGCGCGCT 2164
Qy 1500 TGGCGTAAATCATGGTCAATAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCAAAATCCAC 1559
Db 2163 TGGCGTAAATCATGGTCAATAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCAAAATCCAC 2104
Qy 1560 ACAACATACGAGCCGGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATAGTGTAGCTAAC 1619
Db 2103 ACAACATACGAGCCGGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATAGTGTAGCTAAC 2044
Qy 1620 TCACATTAATTTGCGTTGCGTCACTGCGCTTTTCCAGTCGGGAAACCTGTCGTCGACG 1679
Db 2043 TCACATTAATTTGCGTTGCGTCACTGCGCGCTTTTCCAGTCGGGAAACCTGTCGTCGACG 1984
Qy 1680 TGCATTAATGAATCGGCCAAACCGCGGGGAGAGCGGTTTGGTATTGGGGGCTCTTCCG 1739
Db 1983 TGCATTAATGAATCGGCCAAACCGCGGGGAGAGCGGTTTGGTATTGGGGGCTCTTCCG 1924
Qy 1740 CTTCTCGCTCATGATCTGCTGCGCTCGGTGTTCCGCTGGCGGAGCGGTATCAGCTC 1799
Db 1923 CTTCTCGCTCATGATCTGCTGCGCTCGGTGTTCCGCTGGCGGAGCGGTATCAGCTC 1864
Qy 1800 ACTCAAGGCGGTAAATCGGTTATCCACAGAAATCAGGGGATAACGACGAAAGAACATGT 1859
Db 1863 ACTCAAGGCGGTAAATCGGTTATCCACAGAAATCAGGGGATAACGACGAAAGAACATGT 1804
Qy 1860 GAGCAAAAGGCCAGCAAAAGGCCAGGAACCGGTAAAGGCCGCGTGTGCTGGCGTTTTTCC 1919
Db 1803 GAGCAAAAGGCCAGCAAAAGGCCAGGAACCGGTAAAGGCCGCGTGTGCTGGCGTTTTTCC 1744
Qy 1920 ATAGGCTCGGCCCTCGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGAA 1979
Db 1743 ATAGGCTCGGCCCTCGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGAA 1684

Qy 1980 ACCCGACAGGACTATAAGATACCAAGCGTTTCCCTCGGAAGCTCCCTCGTGGCTCTC 2039
Db 1683 ACCCGACAGGACTATAAGATACCAAGCGTTTCCCTCGGAAGCTCCCTCGTGGCTCTC 1624
Qy 2040 CTGTTCCGACCTCGCGCTTACCGGATACCTGTGCGCTTCTCCCTCGGAAAGCGTGG 2099
Db 1623 CTGTTCCGACCTCGCGCTTACCGGATACCTGTGCGCTTCTCCCTCGGAAAGCGTGG 1564
Qy 2100 CGCTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCCGTTAGGTGAGTTCGCTCCAAGC 2159
Db 1563 CGCTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCCGTTAGGTGAGTTCGCTCCAAGC 1504
Qy 2160 TGGGCTGTGTGCAGAAACCCCGCTTACGCCGCGCTTTCAGCGCTTATTCGCGTAACTATC 2219
Db 1503 TGGGCTGTGTGCAGAAACCCCGCTTACGCCGCGCTTTCAGCGCTTATTCGCGTAACTATC 1444
Qy 2220 GTCTTGAAGTCAACCCCGTAAAGACAGACTTATGCCACTGGCAGACCACTGGTAACA 2279
Db 1443 GTCTTGAAGTCAACCCCGTAAAGACAGACTTATGCCACTGGCAGACCACTGGTAACA 1384
Qy 2280 GGATTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACT 2339
Db 1383 GGATTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACT 1324
Qy 2340 ACGGCTACACTAGAAGGACAGTATTTTGGTATCTGCGCTCTGCTGAAGCAGTTACTTCG 2399
Db 1323 ACGGCTACACTAGAAGGACAGTATTTTGGTATCTGCGCTCTGCTGAAGCAGTTACTTCG 1264
Qy 2400 GAAAAAGATTGGTAGTCTTTGATCCCGCAACAAACCAACCGCTGGTAGCGGTGTTTTT 2459
Db 1263 GAAAAAGATTGGTAGTCTTTGATCCCGCAACAAACCAACCGCTGGTAGCGGTGTTTTT 1204
Qy 2460 TTGTTTGCACGACGAGATTACGCGCAGAAAAAAGGATCTCAGAAAGTCTTTTGAATCT 2519
Db 1203 TTGTTTGCACGACGAGATTACGCGCAGAAAAAAGGATCTCAGAAAGTCTTTTGAATCT 1144
Qy 2520 TTTCTACGCGGTCTGACGCTCAGTGGAAACGAAACTCAGTTAAGGGATTTTGGTCAAGA 2579
Db 1143 TTTCTACGCGGTCTGACGCTCAGTGGAAACGAAACTCAGTTAAGGGATTTTGGTCAAGA 1084
Qy 2580 GATTATCAAAAAGGATCTTCACTAGATCTTTTAAATTAATAATGAAGTTTAAATCAA 2639
Db 1083 GATTATCAAAAAGGATCTTCACTAGATCTTTTAAATTAATAATGAAGTTTAAATCAA 1024
Qy 2640 TCTAAAGTATATACGATTAACCTTGGTCTGACAGTTACCAATGCTTAATCAGTAGGCGAC 2699
Db 1023 TCTAAAGTATATACGATTAACCTTGGTCTGACAGTTACCAATGCTTAATCAGTAGGCGAC 964
Qy 2700 CTATCTCAGCGATCTGTCTATTTTGGTTTCCATCATAGTTTGCCTGACTCCCGCTCGTGTAGA 2759
Db 963 CTATCTCAGCGATCTGTCTATTTTGGTTTCCATCATAGTTTGCCTGACTCCCGCTCGTGTAGA 904
Qy 2760 TAACTACGATACGGAGGGCTTACCATCTGCGCCAGTGTGCAATGATACCGCGAGACC 2819
Db 903 TAACTACGATACGGAGGGCTTACCATCTGCGCCAGTGTGCAATGATACCGCGAGACC 844
Qy 2820 CACGCTCACCGGCTCCAGATTTATCAGCAATAAACCCAGCCAGCGGAGGCGGAGCGCA 2879
Db 843 CACGCTCACCGGCTCCAGATTTATCAGCAATAAACCCAGCCAGCGGAGGCGGAGCGCA 784
Qy 2880 GAAAGTGTCTGCAACTTTATTCGCTCCATCCAGTCTATTAAATTTGTTGCGGGGAAGCTA 2939
Db 783 GAAAGTGTCTGCAACTTTATTCGCTCCATCCAGTCTATTAAATTTGTTGCGGGGAAGCTA 724
Qy 2940 GAGTAAAGTAGTTCGCCAGTTAATAGTTTTCGCAACAGTTTGGTTCGCAATTCGTCACAGCATCG 2999
Db 723 GAGTAAAGTAGTTCGCCAGTTAATAGTTTTCGCAACAGTTTGGTTCGCAATTCGTCACAGCATCG 664
Qy 3000 TGGTGTACGCTCGCTCGTTTGGTATGGCTTCAATTCAGCTCCGTTCCCAACGATCAAGC 3059
Db 663 TGGTGTACGCTCGCTCGTTTGGTATGGCTTCAATTCAGCTCCGTTCCCAACGATCAAGC 604

2579	TCGATACGATTAAACAAGGGTATACACTTTCAAACTTGACTTCAGACACGGCTCTGTGTAGTTC	2520
1141	AAGCGGCATGGAAAGAGTGTGGCCGAGGATGACTCTGCTGTCATTTGACGTTGACGCGAAAC	1200
2519	CCGTCATCTTTGAAAGATATAGTGCCTGCTGTATACATAAACCTTCGGGCATGGCACTCTTG	2460
1201	GCAGCTTTACCATGATGATTCGGGAAAGGTGTGGCCATGCGAGCCTTTTAAACGGTGAACTGT	1360
2459	AAAAAGTCATGCCCTTTTCATATGATCCGGATAACCGGAAAAGCAATTGAAACACCATTAAGAG	2400
1261	TCGTTTCAGGCCACTCGGATACCACTTCGTTCGGCTTTTCCGACACACAGTTCGGATGG	1320
2399	AAAGTAGTGCAAGTGTGTGGCCATGTGGAACAGGTAGTTTTCCAGTAGTGCAAAATAAAATTTA	2340
1321	TCAGCCGGAAGCGCATCAGCAACCCGAAACAATACCGCGACAGCCGAACTGCCGTGCCG	1380
2339	AGGTTAGCTTTTCGTATGTAGCATCACTTTCACCTCTCCACTGACAGAAAATTTGTGC	2280
1381	GTGTGCAGATTAAATGACAGCGGTGCGCGCTGGGATATTACGTACAGCAGGACGGGTATC	1440
2279	CCATTAAACATCAACCATCTAAATTCAACAAGAATTGGGACAACTCCAGTGAAGAATGTTCTCT	2220
1441	CTGCTCGATGCCGACAGAAATGGACATGGATACCCCGTGATGTTACCCGGCGGGCGCGCTT	1500
2219	C---CTTTACTCATCCGTACCCAGCTTTTGTTCCTTTAGTGAGGGTTAAATGCGCGCTT	2163
1501	GGCGTAATCATGGTCATAGCTGTTTTCTGTGTGAAATTTGTTATCCGCTCACAATTCACACA	1560
2162	GGCGTAATCATGGTCATAGCTGTTTTCTGTGTGAAATTTGTTATCCGCTCACAATTCACACA	2103
1561	CAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCCTAATGAGTGAGCTAACT	1620
2102	CAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCCTAATGAGTGAGCTAACT	2043
1621	CACATTAATTGCGTTGCGCTCACTGCCGCTTTCAGTCGGGAAAACCTGTGTCGACGT	1680
2042	CACATTAATTGCGTTGCGCTCACTGCCGCTTTCAGTCGGGAAAACCTGTGTCGACGT	1983
1681	GCATTAATGAATCGGCCAAACGCGGGGAGAGCGGTTTCGTTATTCGGCGCTCTTCCGC	1740
1982	GCATTAATGAATCGGCCAAACGCGGGGAGAGCGGTTTCGTTATTCGGCGCTCTTCCGC	1923
1741	TTCTCTCGCTCACTGACTCGCTGCGCTCGGTCGTTTCGCTCGCGGACGCGGTATCAGCTCA	1800
1922	TTCTCTCGCTCACTGACTCGCTGCGCTCGGTCGTTTCGCTCGCGGACGCGGTATCAGCTCA	1863
1801	CTCAAGGCGGTAAATACGGTTATCCACAGAAATCAAGGGATAAACGAGGAAGAACATGTG	1860
1862	CTCAAGGCGGTAAATACGGTTATCCACAGAAATCAAGGGATAAACGAGGAAGAACATGTG	1803
1861	AGCAAAAGGCCACCAAAAGCCAGAACCTTAAAGCCGCGTGTCTGCGGTTTTTCCA	1920
1802	AGCAAAAGGCCACCAAAAGCCAGAACCCGTTAAAGGCCGCGTGTCTGCGGTTTTTCCA	1743
1921	TAGGCTCCGCCCTTCAGCAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAA	1980
1742	TAGGCTCCGCCCTTCAGCAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAA	1683
1981	CCGACAGGACTATAAAGATACAGCGGTTTTCCCTTCGGAAGCTCCTCTGTGCGCTCTCC	2040
1682	CCGACAGGACTATAAAGATACAGCGGTTTTCCCTTCGGAAGCTCCTCTGTGCGCTCTCC	1623
2041	TGTTCCGACCTGCGCTTACCGGATACCTGTCCGCTTTTCTCCCTTCGGAAGCTGGC	2100
1622	TGTTCCGACCTGCGCTTACCGGATACCTGTCCGCTTTTCTCCCTTCGGAAGCTGGC	1563
2101	GCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTTCGGTGTAGGTCTGTCGCTCCAAGCT	2160
1562	GCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTTCGGTGTAGGTCTGTCGCTCCAAGCT	1503
2161	GGGCTGTGTCAAGAAACCCCGCTTCAGCCGACCGCTGCGCTTATTCGGTAACTATCG	2220
1502	GGGCTGTGTCAAGAAACCCCGCTTCAGCCGACCGCTGCGCTTATTCGGTAACTATCG	1443

Qy	2221	TCTTGA	GTCG	CAACCCCGG	TAAAGACAGACG	ACTTATCGCC	ACTGCGC	AGAGCC	ACTGGT	AAACAG	2280
Db	1442	TCTTGA	GTCG	CAACCCCGG	TAAAGACAGACG	ACTTATCGCC	ACTGCGC	AGAGCC	ACTGGT	AAACAG	1383
Qy	2281	GATTAG	CAGAGC	GAGGTATG	TAGCGCGT	GTCTACAGAGT	CTCTCGAAGT	TGGTGGCT	TAACTA	CTA	2340
Db	1382	GATTAG	CAGAGC	GAGGTATG	TAGCGCGT	GTCTACAGAGT	CTCTCGAAGT	TGGTGGCT	TAACTA	CTA	1323
Qy	2341	CGGCTA	CAC	TAGAAG	GACAGTATTTGG	TATCTCGCGT	CTCTGCTGAAGC	CAAGT	TACCTTCG	G	2400
Db	1322	CGGCTA	CAC	TAGAAG	GACAGTATTTGG	TATCTCGCGT	CTCTGCTGAAGC	CAAGT	TACCTTCG	G	1263
Qy	2401	AAAAAG	ATGGT	TAGTCTT	GATCTCGGCA	AAACAAACACCG	CTGTGTAG	CGGTGGT	TTTT		2460
Db	1262	AAAAAG	ATGGT	TAGTCTT	GATCTCGGCA	AAACAAACACCG	CTGTGTAG	CGGTGGT	TTTT		1203
Qy	2461	TGTTTGA	CAGCAG	CAGATTAACGG	CGACAGAAAAAG	GATCTCAAGAA	GATCTCTTTGAT	CTT			2520
Db	1202	TGTTTGA	CAGCAG	CAGATTAACGG	CGACAGAAAAAG	GATCTCAAGAA	GATCTCTTTGAT	CTT			1143
Qy	2521	TTCTA	CGGGTCT	GACGCTCAGT	GGAACGAAAACT	CACGTTAAGG	GAATTTTGGT	CTATGAG			2580
Db	1142	TTCTA	CGGGTCT	GACGCTCAGT	GGAACGAAAACT	CACGTTAAGG	GAATTTTGGT	CTATGAG			1083
Qy	2581	ATTATCA	AAAGAGATCTT	CACCTAGATCC	TTTAAATTTAAAT	TAAATGAAGT	TTTTTAAATCAAT				2640
Db	1082	ATTATCA	AAAGAGATCTT	CACCTAGATCC	TTTAAATTTAAAT	TAAATGAAGT	TTTTTAAATCAAT				1023
Qy	2641	CTAAG	TATATAG	AGTAAACTTTGGT	CTGACAGTTACCA	TGCTTAATCAGT	GAGGCACC				2700
Db	1022	CTAAG	TATATAG	AGTAAACTTTGGT	CTGACAGTTACCA	TGCTTAATCAGT	GAGGCACC				963
Qy	2701	TATCTCAG	CGATCTG	CTATTTGGTTCAT	CCATAGTTCCTG	ACTCCCGT	CGTGTAGAT				2760
Db	962	TATCTCAG	CGATCTG	CTATTTGGTTCAT	CCATAGTTCCTG	ACTCCCGT	CGTGTAGAT				903
Qy	2761	AACTAC	GATACGGG	AGGCTTACCAT	CTGCGCCAGTGT	GCAATGATACCGG	GAGACCC				2820
Db	902	AACTAC	GATACGGG	AGGCTTACCAT	CTGCGCCAGTGT	GCAATGATACCGG	GAGACCC				843
Qy	2821	ACGCTCA	CCGGCTCCAGAT	TTATCAGCA	ATAACACCGC	AGCCGGAAGG	CGCAGCCAG				2880
Db	842	ACGCTCA	CCGGCTCCAGAT	TTATCAGCA	ATAACACCGC	AGCCGGAAGG	CGCAGCCAG				783
Qy	2881	AAGTGGT	CTCTGCAACTTTAT	CCGCTCCATC	CGAGTCTATTAAT	TGTTGCGG	GAAGCTAG				2940
Db	782	AAGTGGT	CTCTGCAACTTTAT	CCGCTCCATC	CGAGTCTATTAAT	TGTTGCGG	GAAGCTAG				723
Qy	2941	AGTAAGT	AGTTCG	CGAGTTAATAGTT	TGTCGCAACGTT	TGTTGCCA	TATGCTACAGGCAT	CGT			3000
Db	722	AGTAAGT	AGTTCG	CGAGTTAATAGTT	TGTCGCAACGTT	TGTTGCCA	TATGCTACAGGCAT	CGT			663
Qy	3001	GGTGTCA	CGCTCGT	CTGTTGGTAT	GCTTCATTCAGCT	CCGGTTC	CCAAAGTCAAGG	CGC			3060
Db	662	GGTGTCA	CGCTCGT	CTGTTGGTAT	GCTTCATTCAGCT	CCGGTTC	CCAAAGTCAAGG	CGC			603
Qy	3061	AGTTACA	TGATCCCC	ATGTTGTGCA	AAAAAGCGGTTAG	CTCTCGGT	CCTCCGAT	CGT			3120
Db	602	AGTTACA	TGATCCCC	ATGTTGTGCA	AAAAAGCGGTTAG	CTCTCGGT	CCTCCGAT	CGT			543
Qy	3121	TGTCAGA	AGTAAGTTG	CGCCAGTGTTAT	CACATCATG	TTTATGCGCAG	CACTGATTAATTC				3180
Db	542	TGTCAGA	AGTAAGTTG	CGCCAGTGTTAT	CACATCATG	TTTATGCGCAG	CACTGATTAATTC				483
Qy	3181	TCTTACT	GTCTA	TGCGCATCCG	TAAAGTGC	TTTTCTGTG	ACTGTTAGTACTCA	CAAGTC			3240
Db	482	TCTTACT	GTCTA	TGCGCATCCG	TAAAGTGC	TTTTCTGTG	ACTGTTAGTACTCA	CAAGTC			423
Qy	3241	ATTCTGA	AAATAGT	GTATATGCGG	CGACCGAGT	GTGCTTTG	CCCGCGTCAATA	CGGATAA			3300
Db	422	ATTCTGA	AAATAGT	GTATATGCGG	CGACCGAGT	GTGCTTTG	CCCGCGTCAATA	CGGATAA			363

```
QY 3301 TACCGCGCCACATAGCAGAACTTTAAAGTGCTCATCATTTGAAAACGTTCTTCGGGGCG 3360
Db 362 TACCGCGCCACATAGCAGAACTTTAAAGTGCTCATCATTTGAAAACGTTCTTCGGGGCG 303
QY 3361 AAAACTCTCAAGATCTTACCGCTGTGAGATCCAGTTCGATGTAAACCACTCGTGCCAC 3420
Db 302 AAAACTCTCAAGATCTTACCGCTGTGAGATCCAGTTCGATGTAAACCACTCGTGCCAC 243
QY 3421 CAACTGATCTTCAGCATCTTTTACCTTCCACAGCGTTTCTGGGTGAGCAAAAACAGGAAG 3480
Db 242 CAACTGATCTTCAGCATCTTTTACCTTCCACAGCGTTTCTGGGTGAGCAAAAACAGGAAG 183
QY 3481 GCAAAATCCGCAAAAACGGGAATAGCGCGACACGGAAATGTTGAATCTCATACTCTT 3540
Db 182 GCAAAATCCGCAAAAACGGGAATAGCGCGACACGGAAATGTTGAATCTCATACTCTT 123
QY 3541 CTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTCTCATGCGGATACATATT 3600
Db 122 CTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTCTCATGAGCGGATACATATT 63
QY 3601 TGAATGTTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCGGAAAAAGTGCC 3660
Db 62 TGAATGTTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCGGAAAAAGTGCC 3
QY 3661 AC 3662
Db 2 AC 1

RESULT 8
US-10-206-030-4/c
; Sequence 4, Application US/10206030
; Publication No. US20020177160A1
; GENERAL INFORMATION:
; APPLICANT: Padgett, Hal S.
; APPLICANT: Lindbo, John A.
; APPLICANT: Fitzmaurice, Wayne P.
; TITLE OF INVENTION: A Method of Increasing Complementarity
; FILE REFERENCE: In A Heteroduplex
; CURRENT APPLICATION NUMBER: US/10/206,030
; PRIOR FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: US/10/066,390
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 3637
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-10-206-030-4

Query Match 72.8%; Score 2664.8; DB 13; Length 3637;
Best Local Similarity 83.8%; Pred. No. 0;
Matches 3070; Conservative 0; Mismatches 567; Indels 25; Gaps 4;

QY 1 CTAATTTGTAAGCGTTAATTTTGTAAATTCGGTTAAATTTTGTAAATCAGCTC 60
Db 3637 CTAATTTGTAAGCGTTAATTTTGTAAATTCGGTTAAATTTTGTAAATCAGCTC 3578
QY 61 ATTTTAAACCAATAGGCGCAAAATCGCAAAATCCCTTAATAATCAAAAGAAATAGACCGA 120
Db 3577 ATTTTAAACCAATAGGCGCAAAATCGCAAAATCCCTTAATAATCAAAAGAAATAGACCGA 3518
QY 121 GATAGGGTTGAGTGTGTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGAAACGTGACTC 180
Db 3517 GATAGGGTTGAGTGTGTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGAAACGTGACTC 3458
QY 181 CAACTGCAAGGGCGGAAAAACCGTCTATCAGGGCGATGCCCCACTACGTGAACCATCACC 240
Db 3457 CAACTGCAAGGGCGGAAAAACCGTCTATCAGGGCGATGCCCCACTACGTGAACCATCACC 3398
```

```
QY 241 CTAATCAAGTTTTTTGGGTCGAGGTGCGGTAAAGCACTAAATCGGAAACCTTAAAGGAG 300
Db 3397 CTAATCAAGTTTTTTGGGTCGAGGTGCGGTAAAGCACTAAATCGGAAACCTTAAAGGAG 3338
QY 301 CCCCCGATTTAGAGCTTGACGGGGAAGCGCGCAACGTGGCGAAGAAAGGAGGGAAGAA 360
Db 3337 CCCCCGATTTAGAGCTTGACGGGGAAGCGCGCAACGTGGCGAAGAAAGGAGGGAAGAA 3278
QY 361 AGCGAAAGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCAACGCTGCGCTAACCCAC 420
Db 3277 AGCGAAAGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCAACGCTGCGCTAACCCAC 3218
QY 421 CACACCGCGCGCTTAATGCGCGCTACAGGGCGGTCCCAATTCGCCAATTCAGGTGCG 480
Db 3217 CACACCGCGCGCTTAATGCGCGCTACAGGGCGGTCCCAATTCGCCAATTCAGGTGCG 3158
QY 481 CAACTGTTGGGAAGGGCGATCGGTGCGGCGCTTTCGCTATTACGCCAGCTGCGGAAGG 540
Db 3157 CAACTGTTGGGAAGGGCGATCGGTGCGGCGCTTTCGCTATTACGCCAGCTGCGGAAGG 3098
QY 541 GGGATGTGCTCAAGGCGGATTAAGTTGGGTAAACGCCAGCGGTTTCCAGTCACGAGTTG 600
Db 3097 GGGATGTGCTCAAGGCGGATTAAGTTGGGTAAACGCCAGCGGTTTCCAGTCACGAGTTG 3038
QY 601 TAAAAACGACGGCCAGTGAGCGCGCTCGTTCAATTCAGTTTGTGAACCCGTGGAGGACGG 660
Db 3037 TAAAAACGACGGCCAGTGAGCGCGCTCGTTCAATTCAGTTTGTGAACCCGTGGAGTCCA 2978
QY 661 GCAGACTCGCGGTGCAAAATGTTTTCAGCGTGATGGAGCAGATGAAGATGCTCCACAC 720
Db 2977 CCGCGGTGCGCG-----CGGCTCTAGAACTAGTGAGTCCCGGCTCCCG 2939
QY 721 GCTGCAGAACACGCGAGCTAGATTAAACCTTAGAAAGATAATCATATTGTGCACGTCGTTAA 780
Db 2938 GCTGCAGGAATTCATTATTGTATAGTTTCAATGCATGCTAATCCAGCAGCAGTTA 2879
QY 781 AGATAATCATGCGTAAAAATTCAGCATGGATCTGTAAATACGACTCACTATATAGGGCGAAT 840
Db 2878 CAAACTCAAGAAGGACCATGTGTCAGCTTTTTCGTTGGGATCTTTCGAAAAGGCGAGATT 2819
QY 841 TGGGTACCGGGCGCGCGCGCTCGAGGTGCGAGGTATCGATAAGCTTGATATCGAATTCCTGC 900
Db 2818 GTGTCGACAGGTAATGTTGTCTGTGTAAAGACAGCGGCGCATCGCCAAATGGAGTATTTT 2759
QY 901 AGCCCGGGGATCCACTAGTTCTAGAGCGCGCGCACCGCGTGGAGCTCCAGCTTTTGT 960
Db 2758 GTTGATTAATGTTCTGCTAGTTGAAACGATCCATCTTCAATGTTGTGGCGAATTTTGAAGT 2699
QY 961 TCCCTTTTATGAGGGTTAATTAGATCCCATGGGTCAATTTTACGCAAGACTATCTTTCTAG 1020
Db 2698 TAGCTTTGATTCCATTCTTTTGTGTCGCGTGATGATATACATTTGTGTGAG-TTATAG 2640
QY 1021 GGTTAATCTAGCTGATCAGGATCATATCGTCGGGTCTTTTTCGGCTCAGTCATCGCC 1080
Db 2639 TTGTACTCGAGTTTGTGTCGAGAAATGTTTCCACTCTTCTTTAAATCAATACCTTTTAAC 2580
QY 1081 CAACTGCGCTATCTGGGCATCGGGGAGGAAAGCCGCTTTTCCCGGAGGTG 1140
Db 2579 TCGATACGATTAACAGGGTATCACCTTCAAACTTGACTTCAGCAGCGCTCTGTAGTTTC 2520
QY 1141 AAGCGGCATGGAAGAGTTTTCGCGGATGACTGCTGTGCAATTTGACGTTGAGCGGAAAC 1200
Db 2519 CGGTCACTTTTGAAAGATATAGTGCCTTCTGTACATAACCTTCGGGCATGGCAGCTCTTG 2460
QY 1201 GCAGGTTTACATGATGATTCGGGAAGGTGTGGCCATGCAACGCTTTTAACGGTGAATGT 1260
Db 2459 AAAAAAGTCATGCGGTTTTCATATGATCCGATTAACGGGAAAGCATTTGAACACCATTAAG 2400
QY 1261 TCGTTTCAGGCCACCTGGGATACAGTTTCGTCGGGCTTTTCCGGACACAGTTCCCGATG 1320
Db 2399 AAAGTAGTGACAAGTGTGGCCCATGTGAACAGGTAGTTTTCAGTAGTGCAAAATAAATTTA 2340
```


QY 1321 TCAGCCGAAGCGCATCAGCAACCCGAAACAATAFCGCGGACAGCGGAACCTGCCGTGCCG 1380
Db 2339 AGGGTAAGCTTCCGATATGTAGCATCACCTTCCACCTCTCCACTGACAGAAAATTGTGCG 2280
QY 1391 GTGTGCAGATTAA TGACAGCGGTGCGGCGCTGGGATATTAGTTCAGCGAGGACGGGTATC 1440
Db 2279 CCAATTAACATCACCATCTAAATTCAACAAGAAATGGGACAACTCCAGTGAAGAAATCTTTCT 2220
QY 1441 CTGGCTGATCCGCGAGAAATGGACATGATACCCCGTGAGTTACCCGCGGGCGCGTT 1500
Db 2219 C---CTTTACTCATCGGTACCCAGCTTTTGTTCCTTTAGTGAGGGTTAAATGCGCGCTT 2163
QY 1501 GCGGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTTGTTATCGCGTCACAATTTCCACA 1560
Db 2162 GCGGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTTGTTATCGGCTCACAAATTTCCACA 2103
QY 1561 CAACATACGAGCGGGAAGCATAAAGTGTAAAGCCTGGGGTGCGCTTAATGAGTGAGTAACT 1620
Db 2102 CAACATACGAGCGGGAAGCATAAAGTGTAAAGCCTGGGGTGCGCTTAATGAGTGAGTAACT 2043
QY 1621 CACATTAATTTGGTTGGCTCACTGCCCGCTTTCAGTTCGGGAACCTGTGTCGACGT 1680
Db 2042 CACATTAATTTGGTTGGCTCACTGCCCGCTTTCAGTTCGGGAACCTGTGTCGACGT 1983
QY 1681 GCATTAATGAATCGGCCAACCGCGGGGAGAGCGGTTTGGCTATTGGGCGCTCTTCCGC 1740
Db 1982 GCATTAATGAATCGGCCAACCGCGGGGAGAGCGGTTTGGCTATTGGGCGCTCTTCCGC 1923
QY 1741 TTCCTCGCTCACTGACTCGCTCGGTTCGGTTCGGCTGCGGCGAGCGGTATCAGCTCA 1800
Db 1922 TTCCTCGCTCACTGACTCGCTCGGTTCGGTTCGGTTCGGGCGAGCGGTATCAGCTCA 1863
QY 1801 CTCAAAAGCGGTAAATACGGTTATCCACAGAAATCAGGGATTAACGAGGAAAGAACATGTG 1860
Db 1862 CTCAAAAGCGGTAAATACGGTTATCCACAGAAATCAGGGATTAACGAGGAAAGAACATGTG 1803
QY 1861 AGCAAAAGCGCAGCAAAAGGCCAGAACCGTAAAAAGGCCGCTTGCTGGGCTTTTCCCA 1920
Db 1802 AGCAAAAGCGCAGCAAAAGGCCAGAACCGTAAAAAGGCCGCTTGCTGGGCTTTTCCCA 1743
QY 1921 TAGGCTCGGCCCTCTGAGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGCGAAA 1980
Db 1742 TAGGCTCGGCCCTCTGAGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGCGAAA 1683
QY 1981 CCCGACAGGACTATAAAGATACCAAGCGGTTTCCCTCGAAGCTCCCTCGTGGCTCTCC 2040
Db 1682 CCCGACAGGACTATAAAGATACCAAGCGGTTTCCCTCGAAGCTCCCTCGTGGCTCTCC 1623
QY 2041 TGTTCCGACCTGCGCTTACCGGATACCTGTGCGCCTTTCTCCCTTCGGGAAGCGTGCC 2100
Db 1622 TGTTCCGACCTGCGCGCTTACCGGATACCTGTGCGCCTTTCTCCCTTCGGGAAGCGTGCC 1563
QY 2101 GCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTGCGGTAGGTGCTGCTCAGAGCT 2160
Db 1562 GCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTGCGGTAGGTGCTGCTCAGAGCT 1503
QY 2161 GGGCTGTGTGACGAACCCCGCTTCAGCCGACCGCTGCGCCTTATCCCGGTAACTATCG 2220
Db 1502 GGGCTGTGTGACGAACCCCGCTTCAGCCGACCGCTGCGCCTTATCCCGGTAACTATCG 1443
QY 2221 TCTTGAGTCCAAACCGGTAAAGACAGACTTATCGCCACTGGCAGAGCCACTGGTAAACAG 2280
Db 1442 TCTTGAGTCCAAACCGGTAAAGACAGACTTATCGCCACTGGCAGAGCCACTGGTAAACAG 1383
QY 2281 GATTAGCAGAGCGAGGTATGTAGGCGGTGTACAGAGTTCTTGAAGTGGTGGCCTAACTA 2340
Db 1382 GATTAGCAGAGCGAGGTATGTAGGCGGTGTACAGAGTTCTTGAAGTGGTGGCCTAACTA 1323
QY 2341 CGGCTACACTAGAGGAAGTATTGGTATCTGGCTCTGTGAGCCAGTTACCTTCGG 2400
Db 1322 CGGCTACACTAGAGGAAGTATTGGTATCTGGCTCTGTGAGCCAGTTACCTTCGG 1263
QY 2401 AAAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACCCAGCTGGTAGCGGTGGTTTTT 2460

Db 1262 AAAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACCCAGCTGGTAGCGGTGGTTTTT 1203
QY 2461 TGTTTGCAGCAGAGATTACGCCAGAAAAGGATCTCAAGAGATCTTTTCATCTT 2520
Db 1202 TGTTTGCAGCAGAGATTACGCCAGAAAAGGATCTCAAGAGATCTTTTCATCTT 1143
QY 2521 TTCTACGGGGTCTGACGCTCAGTGGAAACAAAACCTCAGCTTAAAGGGAATTTTGGTCATGAG 2580
Db 1142 TTCTACGGGGTCTGACGCTCAGTGGAAACAAAACCTCAGCTTAAAGGGAATTTTGGTCATGAG 1083
QY 2581 ATTATCAAAAAGGATCTTCACTAGATCTCTTTTAAATTAATAAATGAAGTTTAAATCAAT 2640
Db 1082 ATTATCAAAAAGGATCTTCACTAGATCTCTTTTAAATTAATAAATGAAGTTTAAATCAAT 1023
QY 2641 CTTAAGTATATATCAGTAAACCTTGGTCTGACAGTTACCAATGCCTTAATCAGTGAGGCACC 2700
Db 1022 CTTAAGTATATATCAGTAAACCTTGGTCTGACAGTTACCAATGCCTTAATCAGTGAGGCACC 963
QY 2701 TATCTCAGCGATCTGTCTATTATTCCTATTCATATCCTAGTTGCCTGACTCCCGCTGCTAGAT 2760
Db 962 TATCTCAGCGATCTGTCTATTATTCCTATTCCTATTCCTAGTTGCCTGACTCCCGCTGCTAGAT 903
QY 2761 AACTACGATACGGAGGGCTTACCATCTGGGCCCGAGTCTGCAATGATACCGGAGACCC 2820
Db 902 AACTACGATACGGAGGGCTTACCATCTGGGCCCGAGTCTGCAATGATACCGGAGACCC 843
QY 2821 ACGCTCACCGCTCCAGATTATCAGCAATAAACAAGCAGCCAGCGAAGCGCGAGCGAG 2880
Db 842 ACGCTCACCGCTCCAGATTATCAGCAATAAACAAGCAGCCAGCGAAGCGCGAGCGAG 783
QY 2881 AAGTGGTCTCTCAACTTTATTCGGCTCCATCCAGTCTATTAAATTTGTTGCCGGAAGCTAG 2940
Db 782 AAGTGGTCTCTCAACTTTATTCGGCTCCATCCAGTCTATTAAATTTGTTGCCGGAAGCTAG 723
QY 2941 AGTAAGTAGTTGCCAGTTAATAGTTTGGCGAAAGTTTGGCCATTTGCTACAGGCATGCT 3000
Db 722 AGTAAGTAGTTGCCAGTTAATAGTTTGGCGAAAGTTTGGCCATTTGCTACAGGCATGCT 663
QY 3001 GGTGTACGCTCGTCTGTTGATGCTTCAATCAGCTCCGTTCCCAACGATCAAGGCG 3060
Db 662 GGTGTACGCTCGTCTGTTGATGCTTCAATCAGCTCCGTTCCCAACGATCAAGGCG 603
QY 3061 AGTTACATGATCCCCCATTTGTGCAAAAAGCGTTAGCTCTTCGCTCCGATCGT 3120
Db 602 AGTTACATGATCCCCCATTTGTGCAAAAAGCGTTAGCTCTTCGCTCCGATCGT 543
QY 3121 TGTCAAGATGAAGTTGGCGCAGTGTATCACTCATGTTATGGCAGCACTGCAATAATTC 3180
Db 542 TGTCAAGATGAAGTTGGCGCAGTGTATCACTCATGTTATGGCAGCACTGCAATAATTC 483
QY 3181 TCTTACTGTCAATGCCATCCGTAAGATGCTTTTGTGACTGTGAGTACTCAACCAAGTC 3240
Db 482 TCTTACTGTCAATGCCATCCGTAAGATGCTTTTGTGACTGTGAGTACTCAACCAAGTC 423
QY 3241 ATTCTGAGAAATAGTGTATGCGCGACCGAGTTGCTTTTGGCCGGCGTCAATACGGGATAA 3300
Db 422 ATTCTGAGAAATAGTGTATGCGCGACCGAGTTGCTTTTGGCCGGCGTCAATACGGGATAA 363
QY 3301 TACCGCGCCCATAGCAGAACTTTAAAGTGTCTCATCATTTGGAACAACTTCTCGGGGG 3360
Db 362 TACCGCGCCCATAGCAGAACTTTAAAGTGTCTCATCATTTGGAACAACTTCTCGGGGG 303
QY 3361 AAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCCAGTTGATTAACCCACTCGTGACCC 3420
Db 302 AAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTGATTAACCCACTCGTGACCC 243
QY 3421 CAACTGATCTTCAGCATCTTTTACTTTTCCAGCGGTTTCTCGGTGAGCAAAAACAGGAAG 3480
Db 242 CAACTGATCTTCAGCATCTTTTACTTTTACCAGCGGTTTCTCGGTGAGCAAAAACAGGAAG 183
QY 3481 GCAAAATGCCGCAAAAAGGGAATAAGGGCGACAGGAAATGTTGAATACTCATACTCTT 3540

Db 182 GCATAATGCGCAAAAGGGAATAGGGCGGACACGGAAATGTTGAATCTACTCTT 123
Qy 3541 CTTTTCATATTAATGAAGCATTTATCAGGGTTATTCCTCATGAGCGGATACATATT 3600
Db 122 CTTTTCATATTAATGAAGCATTTATCAGGGTTATTCCTCATGAGCGGATACATATT 63
Qy 3601 TGAATGTTATTAAGAAATTAACAAATAGGGTTCCGCGCACATTTCCCGGAAAGTGCC 3660
Db 62 TGAATGTTATTAAGAAATTAACAAATAGGGTTCCGCGCACATTTCCCGGAAAGTGCC 3
Qy 3661 AC 3662
Db 2 AC 1
RESULT 9
US-10-211-079-4/c
; Sequence 4, Application US/10211079
; Publication No. US20030148315A1
; GENERAL INFORMATION:
; APPLICANT: Padgett, Hal S.
; APPLICANT: Vaezhongs, Andrew A.
; APPLICANT: Vojdani, Fakhrieh S.
; APPLICANT: Smith, Mark L.
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding CEL I
; FILE OF INVENTION: Endonuclease and Methods of Use Thereof
; FILE REFERENCE: P-LG 5381
; CURRENT APPLICATION NUMBER: US/10/211,079
; PRIOR FILING DATE: 2002-08-01
; PRIOR APPLICATION NUMBER: US 10/098,155
; PRIOR FILING DATE: 2002-03-14
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 3637
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-10-211-079-4
Query Match 72.8%; Score 2664.8; DB 15; Length 3637;
Best Local Similarity 83.8%; Pred. No. 0;
Matches 3070; Conservative 0; Mismatches 567; Indels 25; Gaps 4;
Qy 1 CTAATTTGTAAGCGTTAATATTTGTTAAATTCGGGTTAAATTTTGTAAATCAGCTC 60
Db 3637 CTAATTTGTAAGCGTTAATATTTGTTAAATTCGGGTTAAATTTTGTAAATCAGCTC 3578
Qy 61 ATTTTAAACCAATAGGCGGAAATCGGCAAAATCCCTTATAATCAAAAGAAATAGACCGA 120
Db 3577 ATTTTAAACCAATAGGCGGAAATCGGCAAAATCCCTTATAATCAAAAGAAATAGACCGA 3518
Qy 121 GATAGGTTGAGTGTTCAGTTTGGAAACAGAGTCACCTATTAAGAAAGTGGACTC 180
Db 3517 GATAGGTTGAGTGTTCAGTTTGGAAACAGAGTCACCTATTAAGAAAGTGGACTC 3458
Qy 181 CAACGTCAAAAGGCGGAAACCGTCTATCAGGCGGATGGCCACTACGTGAACCATCAC 240
Db 3457 CAACGTCAAAAGGCGGAAACCGTCTATCAGGCGGATGGCCACTACGTGAACCATCAC 3398
Qy 241 CTAATCAAGTTTTTGGGTCGAGTGGCGTAAAGCACTAAATTCGGAACCCCTAAAGGGAG 300
Db 3397 CTAATCAAGTTTTTGGGTCGAGTGGCGTAAAGCACTAAATTCGGAACCCCTAAAGGGAG 3338
Qy 301 CCCCAGATTAGAGTTGACGGGGAAAGCCGGGAAACGTTGGCGAGAAAGGAGGAGAA 360
Db 3337 CCCCAGATTAGAGTTGACGGGGAAAGCCGGGAAACGTTGGCGAGAAAGGAGGAGAA 3278
Qy 361 AGCGAAAGGAGCGGCGCTAGGCGGCTGCAAGTGTAGCGGTACGCTCGCGGTAAACAC 420
Db 3277 AGCGAAAGGAGCGGCGCTAGGCGGCTGCAAGTGTAGCGGTACGCTCGCGGTAAACAC 3218

Qy 421 CACACCCCGCGCTTAATGCGCGCGCTACAGGGCGCGTCCCATTCGCCATTTCAGGCTGCG 480
Db 3217 CACACCCCGCGCTTAATGCGCGCGCTACAGGGCGCGTCCCATTCGCCATTTCAGGCTGCG 3158
Qy 481 CAACTGTTGGGAAGGCGGATCGTGCGGCCCTCTTCGCTATTACGCGAGCTGCGGAAAGG 540
Db 3157 CAACTGTTGGGAAGGCGGATCGTGCGGCCCTCTTCGCTATTACGCGAGCTGCGGAAAGG 3098
Qy 541 GGGATGTCGCAAGGCGAATTAAGTTTGGGTAAAGCGCAGGGTTTCCAGTCACGAGTTG 600
Db 3097 GGGATGTCGCAAGGCGAATTAAGTTTGGGTAAAGCGCAGGGTTTCCAGTCACGAGTTG 3038
Qy 601 TAAACGACGCGCAGTCAGCGCGCTCTGTTTTCATTCAGTTTTCGACCCCTGGAGGACCG 660
Db 3037 TAAACGACGCGCAGTCAGCGCGCTCTGTTTTCATTCAGTTTTCGACCCCTGGAGTCCA 2978
Qy 661 GCAGACTCGCGGTGCAAAATGTGTTTACAGCGTGATGGAGCAGATGAAGATGTCGACAC 720
Db 2977 CCGCGTGGCGG-----CGCTCTAGAACTAGTGGATCCCC--CGG 2939
Qy 721 GCTGCAGAACACGCGAGCTAGATTAAACCCCTAGAAAGATAATCATATTGTGACGTCAGTTAA 780
Db 2938 GCTGCAGGAATTTCTTATTGTTATAGTTTCATCCATGCCATGTGTAAATCCCGACGACGATT 2879
Qy 781 AGATAATCATGCGTAAATTTGACGATGGATCTGTAAATACGACTCACTATAGGGCGAAT 840
Db 2878 CAAACTCAAGAAAGGACCATGTGTCACGCTTTTCTGTTGGGATCTTTCGAAAGGCGAGATT 2819
Qy 841 TGGGTACCGGGCCCCCTCGAGGTCGACGGTATCGATAAGCTTGATATCGAATTCCTGC 900
Db 2818 GTGTCACAGGTAATGTTGTCGTAAAGGACAGGGCCATCGCCAAATGGAGTATTTT 2759
Qy 901 AGCCCGGGGGATCCACTAGTTCTTAGAGCGCGCGCCACCGCGGTGGAGCTCCAGCTTTTGT 960
Db 2758 GTTGATAATGTCGTCTAGTTGAAACGATCCATCTTCAATGTTGTGCGCAATTTTGAAGT 2699
Qy 961 TCCCTTTAGTAGGGTTAATAGATCCCATGGTCAATTTTACGAGAGACTATCTTTCTAG 1020
Db 2698 TAGCTTTGATTCCTTTCTGTTGTCGCGGTGATGATACATATGTGTGAG--TTATAG 2640
Qy 1021 GGTAACTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCC 1080
Db 2639 TTGTAATCTCAGTTTGTGTCGAGAAATGTTTTCATCTCTTTAAATCAATACCTTTTAAC 2580
Qy 1081 CAACTGCGCGCTATCTGGGCATCGGGGAGGAAAGCCCGTGCCTTTTCCCGGAGGTG 1140
Db 2579 TCGATACGATTAAAGGGTATCACTTCAAACTTGACTTCAGCAGCGCTTGTAGTTC 2520
Qy 1141 AAGCGGATGGAAGAGTTTGGCGAGGATGACTGTGCTGCAATTGACGTTGACGGAAC 1200
Db 2519 CCGTCATCTTTGAAAGATATAGTGGCTTCTCTGTACATAACCTTCGGGCGATGGCACTTTG 2460
Qy 1201 GCAGTTTACCATGATGATTTCGGGAAGGTGGCCATGACCGCTTTAAACGGTGAACGT 1260
Db 2459 AAAAAAGTCATGCGCTTCAATATGATCCGGATTAACGGGAAAGCAITTGAACACCATAGAG 2400
Qy 1261 TCGTTTCAGCCCACTCGGATACAGTTTCGTGCGGCTTTTTCGGGACACAGTTTCCGGATG 1320
Db 2399 AAGTAGTGACAGTTTGGCCATGGAACAGGTAGTTTTCAGTAGTGCAAAATAATTA 2340
Qy 1321 TCAGCCCGAAGCGCATCAGCAACCCGAAACAATAACCGGAGACAGCGGAACTGCGGCGCG 1380
Db 2339 AGGGTAAGCTTTCCGTTATGTAGCATCACCTTCAACCTCTCCCACTGACAGAAATTTGTGC 2280
Qy 1381 GTGTCAGATTATGACAGCGGTGGCGCTGGGATATTACGTACAGGAGGACGGGTATC 1440
Db 2279 CCATTAAACATACCATCTTAATTCACAGAAATGGGACAACTCCAGTGAAGAGTTCTTCT 2220
Qy 1441 CTGGCTGGATGCGCAGAAATGGACATGATGATACCCCGTCAGTTTACCCGGCGGCGCGCTT 1500
Db 2219 C---CTTACTCATCGGTACCCAGCTTTTGTTCCTTTAGTGAGGGTTAATTCGCGGCTT 2163
Qy 1501 GCGTAATCATGGTCATAGCTGTTTCTGTGTGAAATTTGTTATTCGCTCACAATTCACAC 1560

Db 2162 |||||GGCGTAATCATGGTCATAGCTGTTTCTGTGTGAAATGTTATCGGCTCACAAATCCACA2103
Qy 1561 CAACATACGAGCCGGAAGCATAAAGTGTAAAGCCCTGGGGTGCCTAAATGAGTGAGCTAACT1620
Db 2102 CAACATACGAGCCGGAAGCATAAAGTGTAAAGCCCTGGGGTGCCTAAATGAGTGAGCTAACT2043
Qy 1621 CACATTAATTTGGTTGCGCTCACTGCCCCGCTTTCCAGTTCGGGAAACCTGTGTCGCCAGCT1680
Db 2042 CACATTAATTTGGTTGCGCTCACTGCCCCGCTTTCCAGTTCGGGAAACCTGTGTCGCCAGCT1983
Qy 1681 GCATTAATGAATCGGCCAAACCGCGGGGAGAGAGCGGTTTGGTATTTGGGCGCTCTTCCGC1740
Db 1982 GCATTAATGAATCGGCCAAACCGCGGGGAGAGAGCGGTTTGGTATTTGGGCGCTCTTCCGC1923
Qy 1741 TTCCCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCA1800
Db 1922 TTCCCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCA1863
Qy 1801 CTCAAAAGCGGTAATACGGTTATCCACAGAATCAGGGGATAAACGAGGAAAGAAACATGTG1860
Db 1862 CTCAAAAGCGGTAATACGGTTATCCACAGAATCAGGGGATAAACGAGGAAAGAAACATGTG1803
Qy 1861 AGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAGAGCCGCTTGTGCGGTTTTCCTCA1920
Db 1802 AGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAGAGCCGCTTGTGCGGTTTTCCTCA1743
Qy 1921 TAGGCTCGCCCCCTGACGAGCATCAAAAATCGAGCTCAAGTCAGAGGTGCGGAA1980
Db 1742 TAGGCTCGCCCCCTGACGAGCATCAAAAATCGAGCTCAAGTCAGAGGTGCGGAA1683
Qy 1981 CCCGACGAGACTATAAAGATACACAGGCGTTTCCCTCGGAAGCTCCCTCGCGCTCTCC2040
Db 1682 CCCGACGAGACTATAAAGATACACAGGCGTTTCCCTCGGAAGCTCCCTCGCGCTCTCC1623
Qy 2041 TGTTCCGACCCCTGCGCTTACCGGATACCTGTGCGCTTTTCTCCCTCGGAAGCGTGCC2100
Db 1622 TGTTCCGACCCCTGCGCTTACCGGATACCTGTGCGCTTTTCTCCCTCGGAAGCGTGCC1563
Qy 2101 GCTTCTCATAGCTCAGCTGTAGGTATCTCAGTTGCGGTAGGTGCTTCCCTCCAGCT2160
Db 1562 GCTTCTCATAGCTCAGCTGTAGGTATCTCAGTTGCGGTAGGTGCTTCCCTCCAGCT1503
Qy 2161 GGGCTGTGTGACGAACCCCGCTTACGCCGACCGCTGCGCTTATCCGGTAACTATCG2220
Db 1502 GGGCTGTGTGACGAACCCCGCTTACGCCGACCGCTGCGCTTATCCGGTAACTATCG1443
Qy 2221 TCTTGAGTCCAAACCCGGTAAACACGACTTATCGCACTTGGGACGAGCCACTGGTAAACAG2280
Db 1442 TCTTGAGTCCAAACCCGGTAAACACGACTTATCGCACTTGGGACGAGCCACTGGTAAACAG1383
Qy 2281 GATTAGCAGCGAGGTATGTAGGCGGTCTACAGAGTTCTTGAGTGGTGGCTTAACCTA2340
Db 1382 GATTAGCAGCGAGGTATGTAGGCGGTCTACAGAGTTCTTGAGTGGTGGCTTAACCTA1323
Qy 2341 CGGCTACACTAGAAGGACAGTATTGGTATCTGCGCTCTGCTGAAGCCAGTATACCTTCGG2400
Db 1322 CGGCTACACTAGAAGGACAGTATTGGTATCTGCGCTCTGCTGAAGCCAGTATACCTTCGG1263
Qy 2401 AAAAAGAGTTGGTACCTTGATCCGGCAAAACAAACCAACCGCTGGTAGCGGTGTTTTT2460
Db 1262 AAAAAGAGTTGGTACCTTGATCCGGCAAAACAAACCAACCGCTGGTAGCGGTGTTTTT1203
Qy 2461 TGTTTGCAGCAGCATTAACCGCAGAAAAAAGGATCTCAAGAGATCTTTGATCTT2520
Db 1202 TGTTTGCAGCAGCATTAACCGCAGAAAAAAGGATCTCAAGAGATCTTTGATCTT1143
Qy 2521 TTCTACGGGGTCTGACGCTCAGTGGAACGAAACTCAGCTTAAAGGATTTTGGTCAATGAG2580
Db 1142 TTCTACGGGGTCTGACGCTCAGTGGAACGAAACTCAGCTTAAAGGATTTTGGTCAATGAG1083
Qy 2581 ATTATCAAAAAAGGATCTTCACTAGATCCTTTTAAATTAATAATGAAGTTTAAATCAAT2640
|||||

Db 1082 ATTATCAAAAAAGGATCTTCACTAGATCCTTTTAAATTAATAATGAAGTTTAAATCAAT1023
Qy 2641 CTTAAGTATATATGAGTAAACTTGGTCTGACAGTTTACCAATGCTTAAATCAGTGAGGCACC2700
Db 1022 CTTAAGTATATATGAGTAAACTTGGTCTGACAGTTTACCAATGCTTAAATCAGTGAGGCACC963
Qy 2701 TATCTCAGCGATCTGTCTATTTCGTTTCATTCATAGTTGCTTGAATCCCGCTGCTGTAGAT2760
Db 962 TATCTCAGCGATCTGTCTATTTCGTTTCATTCATAGTTGCTTGAATCCCGCTGCTGTAGAT903
Qy 2761 AACTACGATACGGGAGGCTTACCATCTGGCCCCAGTCTGCAATGATACCGGAGACCC2820
Db 902 AACTACGATACGGGAGGCTTACCATCTGGCCCCAGTCTGCAATGATACCGGAGACCC843
Qy 2821 ACGCTCACCGCTCCAGATTTATCAGCAATAAACCAGCAGCCGGAAGGCCGAGCGCAG2880
Db 842 ACGCTCACCGCTCCAGATTTATCAGCAATAAACCAGCAGCCGGAAGGCCGAGCGCAG783
Qy 2881 AAGTGGTCTCGCAACTTTTATCCGCTCCATCCAGTCTATTAAATTTGTTCCCGGAAAGCTAG2940
Db 782 AAGTGGTCTCGCAACTTTTATCCGCTCCATCCAGTCTATTAAATTTGTTCCCGGAAAGCTAG723
Qy 2941 AGTAAAGTAGTTCCGCGATTAATGTTTCCGCAAGCTTGGCCATTTGCTACAGGCATCGT3000
Db 722 AGTAAAGTAGTTCCGCGATTAATGTTTCCGCAAGCTTGGCCATTTGCTACAGGCATCGT663
Qy 3001 GGTGTACGCTCGTCTGTTGGTATGCGTTCATTACAGTCCGTTCCCAACGATCAAGGCG3060
Db 662 GGTGTACGCTCGTCTGTTGGTATGCGTTCATTACAGTCCGTTCCCAACGATCAAGGCG603
Qy 3061 AGTTACATGATPCCCCATGTTGTGCAAAAAAGCGGTAGTCTCTTCCGTCCTCCGATCGT3120
Db 602 AGTTACATGATPCCCCATGTTGTGCAAAAAAGCGGTAGTCTCTTCCGTCCTCCGATCGT543
Qy 3121 TGTCAAGAGTAGTTGGCGCAGTGTATTCACTCATGTTATGGCAGCACTGCTAAATTC3180
Db 542 TGTCAAGAGTAGTTGGCGCAGTGTATTCACTCATGTTATGGCAGCACTGCTAAATTC483
Qy 3181 TCTTACTGTCTATGTCATCCGTAAGATGCTTTCTGTGATCTGTTAGTACTCAACCAAGTC3240
Db 482 TCTTACTGTCTATGTCATCCGTAAGATGCTTTCTGTGATCTGTTAGTACTCAACCAAGTC423
Qy 3241 ATTCTGAGAAATGATGTATGCGGCAACGAGTTGCTTTGCGCGGCTCAATACGGGATAA3300
Db 422 ATTCTGAGAAATGATGTATGCGGCAACGAGTTGCTTTGCGCGGCTCAATACGGGATAA363
Qy 3301 TACCGCGCCATAGCAGAACTTTAAAGTCTCATCATTTGGAACCTTCTTCCGGGCG3360
Db 362 TACCGCGCCATAGCAGAACTTTAAAGTCTCATCATTTGGAACCTTCTTCCGGGCG303
Qy 3361 AAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATTAACCCACTCTGTCGACC3420
Db 302 AAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATTAACCCACTCTGTCGACC243
Qy 3421 CAACTGATCTTCAGCACTCTTTTACCTTTTACCGAGGTTTCTGGGTGAGCAAAAAGGAG3480
Db 242 CAACTGATCTTCAGCACTCTTTTACCTTTTACCGAGGTTTCTGGGTGAGCAAAAAGGAG183
Qy 3481 GCAAAATGCCCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAAATCTCATACTCTT3540
Db 182 GCAAAATGCCCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAAATCTCATACTCTT123
Qy 3541 CCTTTTCAATATTATTGAAGCACTTTATCAGGTTTATTGTCTCATGAGCGGATACATATT3600
Db 122 CCTTTTCAATATTATTGAAGCACTTTATCAGGTTTATTGTCTCATGAGCGGATACATATT63
Qy 3601 TGAATGTATTTAGAAAAATAAAACAAATAGGGGTTCCGCGCACTTTCCCGGAAAGTGCC3660
Db 62 TGAATGTATTTAGAAAAATAAAACAAATAGGGGTTCCGCGCACTTTCCCGGAAAGTGCC3
Qy 3661 AC3662
|||
Db 2AC1

RESULT 10

US-10-356-708-17/c
; Sequence 17, Application US/10356708
; Publication No. US20030157682A1
; GENERAL INFORMATION:
; APPLICANT: Large Scale Biology Corporation
; TITLE OF INVENTION: MISMATCH ENDONUCLEASES AND METHODS OF USE THEREOF
; FILE REFERENCE: P-LG 10100
; CURRENT APPLICATION NUMBER: US/10/356,708
; CURRENT FILING DATE: 2003-01-31
; PRIOR APPLICATION NUMBER: 60/353,722
; PRIOR FILING DATE: 2000-02-01
; PRIOR APPLICATION NUMBER: 10/098,155
; PRIOR FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 10/211,079
; PRIOR FILING DATE: 2002-08-01
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 3637
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Encodes cycle 3 GFP
US-10-356-708-17

Query Match 72.8%; Score 2664.8; DB 16; Length 3637;
Best Local Similarity 83.8%; Pred. No. 0;
Matches 3070; Conservative 0; Mismatches 567; Indels 25; Gaps 4;

QY 1 CTAATTTGAAGCGTTAAATTTTGTAAATTCGGGTTAAATTTTGTAAATTCAGCTC 60
DB 3637 CTAATTTGAAGCGTTAAATTTTGTAAATTCGGGTTAAATTTTGTAAATTCAGCTC 3578

QY 61 ATTTTAAACCAATAGCGCGAAATCGGCAAAATCCCTTATAATCAAAAGATAGACCGA 120
DB 3577 ATTTTAAACCAATAGCGCGAAATCGGCAAAATCCCTTATAATCAAAAGATAGACCGA 3518

QY 121 GATAGGTTGAGTGTGTTTCCAGTTTGGAAACAGAGTCCACTATTAAAGAAAGTGGACTC 180
DB 3517 GATAGGTTGAGTGTGTTTCCAGTTTGGAAACAGAGTCCACTATTAAAGAAAGTGGACTC 3458

QY 181 CAAAGTCAAAAGGCGGCAAAACCGTCTATCAGGCGATGGCCCACTACGTGAACCATCACC 240
DB 3457 CAAAGTCAAAAGGCGGCAAAACCGTCTATCAGGCGATGGCCCACTACGTGAACCATCACC 3398

QY 241 CTAATCAAGTTTTCGGGTCGAGGTGCGGTAAAGCACTAAATCGGAACCCCTAAAGGAG 300
DB 3397 CTAATCAAGTTTTCGGGTCGAGGTGCGGTAAAGCACTAAATCGGAACCCCTAAAGGAG 3338

QY 301 CCCCGATTAGAGCTTGACGGGGAAGCCGGCAACGTCGGCGAGAAAGGAAGAA 360
DB 3337 CCCCGATTAGAGCTTGACGGGGAAGCCGGCAACGTCGGCGAGAAAGGAAGAA 3278

QY 361 AGCGAAGGAGCGGCGCTAGGCGCTGGCAAGTGTAGCGGTACACGCTCGCGTAAACAC 420
DB 3277 AGCGAAGGAGCGGCGCTAGGCGCTGGCAAGTGTAGCGGTACACGCTCGCGTAAACAC 3218

QY 421 CACACCCGCGGCTTAATTCGCGCGCTACAGGCGGCTCCCATTCGCCATTCAGGCTCG 480
DB 3217 CACACCCGCGGCTTAATTCGCGCGCTACAGGCGGCTCCCATTCGCCATTCAGGCTCG 3158

QY 481 CAACTGTGGGAGGCGGATCGGTGGGCTCTTCGCTATTACGCCAGCTCGCGAAGG 540
DB 3157 CAACTGTGGGAGGCGGATCGGTGGGCTCTTCGCTATTACGCCAGCTCGCGAAGG 3098

QY 541 GGGATGTGTCGAAGGCGATTAAAGTTGGGTAAAGCCAGGCTTTCCAGTCAGCAGCTTG 600
DB 3097 GGGATGTGTCGAAGGCGATTAAAGTTGGGTAAAGCCAGGCTTTCCAGTCAGCAGCTTG 3038

QY 601 TAAACGACGCGCAGTGAGCGGCGCTCGTTTCATTCACGTTTTTGAACCCCGTGGAGACGG 660

DB 3037 TAAACGACGCGCAGTGAGCGCGGTAAATACGACTCACTATATAGGGCGAATTTGAGCTCCA 2978
QY 661 GCAGACTCGCGGTGCAAAATGTGTTTTTACAGCGTGATGGAGCAGATGAAGATGCTCGACAC 720
DB 2977 CCGCGTGGCGG-----CGCTCTAGAACTAGTGGATCCC--CGG 2939
QY 721 GCTGCAGAACGCGCAGCTAGATTAAACCTAGAAAGATATCATATTGTGAGCTACGTTAA 780
DB 2938 GCTGCAGGAATCTTATTGTTGATAGTTCATCCATGCCATGTGTAATCCCGACGACGATTA 2879
QY 781 AGATAATCATCGTAAATTTGACGATGGATCTGTAATACGACTCACTATACGGCGAAT 840
DB 2878 CAAACTCAAGAAGGACCATGTGGTACGCTTTCGTTGGGATCTTTCGAAGGGCAGATT 2819
QY 841 TGGGTACCGGGCCCCCTCGAGGTCGACGGTATCGATAAGCTTGATATCGAATTCCTGC 900
DB 2818 GTGTCGACAGGTAAATGCTGCTGTAAGGACAGCGGCATCGCAATTTGGAGTATTT 2759
QY 901 AGCCCGGGGATCCACTAGTTCTAGAGCGGCGCCACCGGGTGGAGCTCCAGCTTTTGT 960
DB 2758 GTTGATAATGCTGCTAGTTGAACGGATCCATCTTCAATGTTGTGGCGAATTTTGAAGT 2699
QY 961 TCCCTTTAGTGAGGGTTAATTAGATCCCATGCGTCAATTTTACGACAGACTATCTTCTAG 1020
DB 2698 TAGCTTTGATTCATTTCTTTGTTGCTGCGGTGATGATATCATTTGTGTAG--TTATAG 2640
QY 1021 GGTAACTTAGCTGCATCAGGATCATGTCGGGTCTTTTTTCCGGCTCAGTCATCGCC 1080
DB 2639 TTGTACTCGAGTTGTGTCGAGAAATGTTTCCATCTCTTTAAATCAATACCTTTTAAAC 2580
QY 1081 CAACTGGCGCTATCTGGGCATCGGGGAGGAGAAAGCCGTCCTTTTCCCGGAGGTTG 1140
DB 2579 TCGATACGATTAAACAGGGTATCACCTTCAAACTTGACTTCAGCAGCGCTTGTGTAGTTC 2520
QY 1141 AAGCGCATGGAAGAGTTTGGCGGAGATGACTGCTGCTGCAATGACGTTGACGGAAAC 1200
DB 2519 CGGTCACTTTGAAAGATATAGTGGCTTCTGTATACATAACCTTCGGGCAATGGCACTCTTG 2460
QY 1201 GCACTGTTTACCATGATGATTCGGGAAGTGTGGCCATGCAACGCTTTTAAAGCTGTA 1260
DB 2459 AAAAGTCATCGCTTTTATATGATCCGGTAAACGGGAAAGCAATTTGAACACCATTAAGAG 2400
QY 1261 TCGTTCAGGCCACCTGGGATACAGTTCTGTCGGCTTTTTCGGGACACAGTTCCGGATGG 1320
DB 2399 AAAGTAGTGACAAGTGTGGCCATGGAAACAGGTAGTTTTCAGTAGTGCATAAATAAATTA 2340
QY 1321 TCAGCCCGAAGCGCATCAGCAACCCGAAACAATACGGCGACAGCCGGAACCTGCGGCGC 1380
DB 2339 AGGTAAGCTTTCCGTTATGATGCACTTCCCTTCCCTCTCCACTGACAGAAATTTGTGC 2280
QY 1381 GTGTGCAGATTAAATGACAGCGGTGCGGCTGGGATATTACGTCAGCGAGGACGGGTATC 1440
DB 2279 CCATTAACTACCATCTAATTTCAACAAGATTGGACAACTCCAGTGAAGTTCTTCT 2220
QY 1441 CTGGCTGGATCGCGCAAAATGGAATGATGATACCCGCTGAGTTACCCGCGGCGCGCTT 1500
DB 2219 C---CTTTACTCATCGTACCCAGCTTTTGTTCCTTTAGTGAGGGTTAATTCGCGCTT 2163
QY 1501 GGGTAATCATGTCATAGCTGTTTCTGTTGAAATTTCTTCCGCTCAGCAATTCACA 1560
DB 2162 GGGTAATCATGTCATAGCTGTTTCTGTTGAAATTTGTTATCCGCTCACAATTCACA 2103
QY 1561 CAACTACAGCGGGAAGCATAAAGTGTAAAGCTGGGCTGCTAATAGTGAAGTAACT 1620
DB 2102 CAACTACAGCGGGAAGCATAAAGTGTAAAGCTGGGCTGCTAATAGTGAAGTAACT 2043
QY 1621 CACTTAATTCGCTTCGCTCACTGCCCCCTTTCCAGTCGGGAAACCTGTCGTCAGCT 1680
DB 2042 CACTTAATTCGCTTCGCTCACTGCCCCCTTTCCAGTCGGGAAACCTGTCGTCAGCT 1983
QY 1681 GCATTAAATCGGCAACGCGCGGAGAGCGGTTTCCGTTATTTGGGGCTCTTCCG 1740

Db 1982 GCATTAATGAATCGGCAACGCGCGGGAGAGGCGGTTTGCGTATTGGCGGCTCTTCGCG 1923
Qy 1741 TTCCCTCGCTCACTGACTCGCTCGGCTCGGTCGCTCGGCTCGGCGGAGCGGATACAGCTCA 1800
Db 1922 TTCCCTCGCTCACTGACTCGCTCGGCTCGGTCGCTCGGCTCGGCGGAGCGGATACAGCTCA 1863
Qy 1801 CTCAAAGCGGTAATACCGTTATCCACAGAAATCAGGGGATACCGAGGAAAGAAATGTTG 1860
Db 1862 CTCAAAGCGGTAATACCGTTATCCACAGAAATCAGGGGATACCGAGGAAAGAAATGTTG 1803
Qy 1861 AGCAAAAGCCAGCAAAAGGCGAGAAACCGTAAAGGCGCGGTTGCTGGCGTTTTTCCA 1920
Db 1802 AGCAAAAGCCAGCAAAAGGCGAGAAACCGTAAAGGCGCGGTTGCTGGCGTTTTTCCA 1743
Qy 1921 TAGGCTCGGCCCCCTGAGAGCATCAAAAATCGACGCTCAAGTTCAGAGTGGCGAAA 1980
Db 1742 TAGGCTCGGCCCCCTGAGAGCATCAAAAATCGACGCTCAAGTTCAGAGTGGCGAAA 1683
Qy 1981 CCGGACAGGACTATAAAGATACCGGCGTTTCCGCTCGAAAGCTCCCTCGTGGCTCTCC 2040
Db 1682 CCGGACAGGACTATAAAGATACCGGCGTTTCCGCTCGAAAGCTCCCTCGTGGCTCTCC 1623
Qy 2041 TGTTCGACCCCTGCGCTTACCGGATACCTGTCGCGCTTTCTCCCTCGGGAAGCGTGGC 2100
Db 1622 TGTTCGACCCCTGCGCTTACCGGATACCTGTCGCGCTTTCTCCCTCGGGAAGCGTGGC 1563
Qy 2101 GCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTTCGGGTAGGTGCTTCGCTCCAAGCT 2160
Db 1562 GCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTTCGGGTAGGTGCTTCGCTCCAAGCT 1503
Qy 2161 GGGCTGTGTGACGAAACCCCGCTTCAGCGCCGACCGCTCGCGCTTATCCGGTAACTATCG 2220
Db 1502 GGGCTGTGTGACGAAACCCCGCTTCAGCGCCGACCGCTCGCGCTTATCCGGTAACTATCG 1443
Qy 2221 TCTTGAGTCCAAACCCGGTAAGACACGACTTATCGCACTGCGAGCAGCACTGTGTAA 2280
Db 1442 TCTTGAGTCCAAACCCGGTAAGACACGACTTATCGCACTGCGAGCAGCACTGTGTAA 1383
Qy 2281 GATTAGCAGAGCGGATGTAGGCGGTCTACAGATTCTGAAGTGGTGGCCCTAACTA 2340
Db 1382 GATTAGCAGAGCGGATGTAGGCGGTCTACAGATTCTGAAGTGGTGGCCCTAACTA 1323
Qy 2341 CGGCTACACTAGAAGCAGATTTTGGTATTTGGCTCTGTGCTGAGCCAGTTTACCTTCGG 2400
Db 1322 CGGCTACACTAGAAGCAGATTTTGGTATTTGGTATCTGCGCTCTGCTGAGCCAGTTTACCTTCGG 1263
Qy 2401 AAAAAGATTGGTAGCTCTTGATCCGGCAACAAACCAACCGCTGTGTAGCGGTGGTTTTT 2460
Db 1262 AAAAAGATTGGTAGCTCTTGATCCGGCAACAAACCAACCGCTGTGTAGCGGTGGTTTTT 1203
Qy 2461 TGTTCGAGCAGCAGATTTACGCGAGAAAGAAAGGATCTCAAGAGATCCTTTGATCTT 2520
Db 1202 TGTTCGAGCAGCAGATTTACGCGAGAAAGAAAGGATCTCAAGAGATCCTTTGATCTT 1143
Qy 2521 TTCTACGGGGTCTGACGCTCAGTGGAAAGAAACCTCAGTTTAAAGGATTTTGGTCAATGAG 2580
Db 1142 TTCTACGGGGTCTGACGCTCAGTGGAAAGAAACCTCAGTTTAAAGGATTTTGGTCAATGAG 1083
Qy 2581 ATTATCAAAAAGGATCTTCACTAGATCCTTTTAAATTAATAAGATTTTAAATCAAT 2640
Db 1082 ATTATCAAAAAGGATCTTCACTAGATCCTTTTAAATTAATAAGATTTTAAATCAAT 1023
Qy 2641 CTAAGATATATAGATTAACCTTGCTGACAGTTTACCAATGCTTAATCAGTGGGACCC 2700
Db 1022 CTAAGATATATAGATTAACCTTGCTGACAGTTTACCAATGCTTAATCAGTGGGACCC 963
Qy 2701 TATCTCAGCGATCTGCTATTTTGGTTTCCATAGTTTGCCTGACTCCCGCGTGGTAGAT 2760
Db 962 TATCTCAGCGATCTGCTATTTGGTTTCCATAGTTTGCCTGACTCCCGCGTGGTAGAT 903
Qy 2761 AACTACGATACGGAGGGCTTACCATCTGGCCCCAGTCTGCAATGATACCGGAGACCC 2820
Db 902 AACTACGATACGGAGGGCTTACCATCTGGCCCCAGTCTGCAATGATACCGGAGACCC 843

RESULT 11

US-10-205-772-4/c
; Sequence 4, Application US/10205772
; Publication No. US20030186261A1
; GENERAL INFORMATION:
; APPLICANT: Padgett, Hal S.
; APPLICANT: Lindbo, John A.
; APPLICANT: Fitzmaurice, Wayne P.
; TITLE OF INVENTION: A Method of Increasing Complementarity
; TITLE OF INVENTION: In A Heteroduplex
; FILE REFERENCE: P-LG 4878

Qy 2821 ACCTCACCGCTCCAGATTTTATCAGCAATAAAACAGCCAGCCGGAAGCGCGAGCGAG 2880
Db 842 ACCTCACCGCTCCAGATTTTATCAGCAATAAAACAGCCAGCCGGAAGCGCGAGCGAG 783
Qy 2881 AAGTGGTCTCTCAACTTTATCCGCTCCATCCAGTCTATAATTTGTTCCGCGAAGCTAG 2940
Db 782 AAGTGGTCTCTCAACTTTATCCGCTCCATCCAGTCTATAATTTGTTCCGCGAAGCTAG 723
Qy 2941 AGTAAGTAGTTGCGCAGTTAATAGTTTCCGCAACGTTTGTTCGCATTTGCTACAGGCATCGT 3000
Db 722 AGTAAGTAGTTGCGCAGTTAATAGTTTCCGCAACGTTTGTTCGCATTTGCTACAGGCATCGT 663
Qy 3001 GGTGTACGCTCGTGGTTTGGTATCGGCTTCAATCAGTCCGCTTCCCAACGATCAAGGCG 3060
Db 662 GGTGTACGCTCGTGGTTTGGTATCGGCTTCAATCAGTCCGCTTCCCAACGATCAAGGCG 603
Qy 3061 AGTTACATGATCCCCCATGTTTGTGCAAAAAAGCGTTAGCTCTCTTCGCTCCGATCGT 3120
Db 602 AGTTACATGATCCCCCATGTTTGTGCAAAAAAGCGTTAGCTCTCTTCGCTCCGATCGT 543
Qy 3121 TGTCAAGATAGTTGGCGCAGTGTATCACTCATGTTTATGGCAGCACTGCTAATAATTC 3180
Db 542 TGTCAAGATAGTTGGCGCAGTGTATCACTCATGTTTATGGCAGCACTGCTAATAATTC 483
Qy 3181 TCTTACTGTCTATGCCATCCGTAAGATGCTTTTCTGTGATCGTGTAGTACTCAACCAAGTC 3240
Db 482 TCTTACTGTCTATGCCATCCGTAAGATGCTTTTCTGTGATCGTGTAGTACTCAACCAAGTC 423
Qy 3241 ATTCTGAGAAATAGTGTATGCGCGCAACGATGTCTTTGCGCGGCGTCAATACGGGATAA 3300
Db 422 ATTCTGAGAAATAGTGTATGCGCGCAACGATGTCTTTGCGCGGCGTCAATACGGGATAA 363
Qy 3301 TACCGCGCACATAGCAGAACTTTTAAAGTGTCTCATCATTTGGTAAACGTTCTTCGGGGCG 3360
Db 362 TACCGCGCACATAGCAGAACTTTTAAAGTGTCTCATCATTTGGTAAACGTTCTTCGGGGCG 303
Qy 3361 AAAACTCTCAAGGATCTTTACCGCTGTTGAGATCCAGTTTCGATGTAAACCCACTCGTGCACC 3420
Db 302 AAAACTCTCAAGGATCTTTACCGCTGTTGAGATCCAGTTTCGATGTAAACCCACTCGTGCACC 243
Qy 3421 CAACGTATCTTCAGCATCTTTTACCTTTTCCACGAGGTTTCTGGGTGAGCAAAAACAGGAAG 3480
Db 242 CAACGTATCTTCAGCATCTTTTACCTTTTCCACGAGGTTTCTGGGTGAGCAAAAACAGGAAG 183
Qy 3481 GCAAAATGCCCAAAAAGGCAATAAGGCGACACGGAATGTTGAAATACTCATACTCTT 3540
Db 182 GCAAAATGCCCAAAAAGGCAATAAGGCGACACGGAATGTTGAAATACTCATACTCTT 123
Qy 3541 CCTTTTTCATATTTATTTGAAGCATTTTATCAGGGTTATTGTCTCATGAGCGGATACATATT 3600
Db 122 CCTTTTTCATATTTATTTGAAGCATTTTATCAGGGTTATTGTCTCATGAGCGGATACATATT 63
Qy 3601 TGAATGTAATTTAGAAAAATAAAATAAGGGGTTCCGCGCACATTTCCCGGAAAAGTCC 3660
Db 62 TGAATGTAATTTAGAAAAATAAAATAAGGGGTTCCGCGCACATTTCCCGGAAAAGTCC 3
Qy 3661 AC 3662
Db 2 AC 1

;; CURRENT APPLICATION NUMBER: US/10/205,772

;; CURRENT FILING DATE: 2002-07-25

;; PRIOR APPLICATION NUMBER: US/10/066,390

;; PRIOR FILING DATE: 2002-02-01

;; NUMBER OF SEQ ID NOS: 15

;; SOFTWARE: FastSeq for Windows Version 4.0

;; SEQ ID NO 4:

;; LENGTH: 3637

;; TYPE: DNA

;; ORGANISM: Artificial Sequence

;; FEATURE:

;; OTHER INFORMATION: synthetic construct

US-10-205-772-4

Query Match 72.8%; Score 2664.8; DB 16; Length 3637;

Best Local Similarity 83.8%; Pred. No. 0;

Matches 3070; Conservative 0; Mismatches 567; Indels 25; Gaps 4;

QY	1	CTAAATTGTAAGCGTTAAATATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTC	60
DB	3637	CTAAATTGTAAGCGTTAAATATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTC	3578
QY	61	ATTTTAAACCAATAGCGCGAAATCGCAAAATCCCTTAAATCAAAAGAAATAGACCGA	120
DB	3577	ATTTTAAACCAATAGCGCGAAATCGCAAAATCCCTTAAATCAAAAGAAATAGACCGA	3518
QY	121	GATAGGTTGAGTTGTTCCAGTTTGGAAACAGAGTCCACTATTAAAGAAAGTGGACTC	180
DB	3517	GATAGGTTGAGTTGTTCCAGTTTGGAAACAGAGTCCACTATTAAAGAAAGTGGACTC	3458
QY	181	CAACGTCAAGGGCGGAAACCGCTCTATCAGGGCGATGCCCACTACGTGAACCATCAC	240
DB	3457	CAACGTCAAGGGCGGAAACCGCTCTATCAGGGCGATGCCCACTACGTGAACCATCAC	3398
QY	241	CTAAATCAAGTTTTCGGGTCGAGGTGCGTAAAGCACTAAATTCGGAACCCCTAAAGGGAG	300
DB	3397	CTAAATCAAGTTTTCGGGTCGAGGTGCGTAAAGCACTAAATTCGGAACCCCTAAAGGGAG	3338
QY	301	CCCCCGATTAGAGCTTGAACGGGAAAGCCGCGAAACGTGGCGAGAAAGGAAGGAA	360
DB	3337	CCCCCGATTAGAGCTTGAACGGGAAAGCCGCGAAACGTGGCGAGAAAGGAAGGAA	3278
QY	361	ACGGAAGAGCGGGCGCTAGGGCGTGCAGTGTAGGGTCAAGCTCGGCTCGGCTTAACAC	420
DB	3277	ACGGAAGAGCGGGCGCTAGGGCGTGCAGTGTAGGGTCAAGCTCGGCTCGGCTTAACAC	3218
QY	421	CACACCGCGCGCTTAATCGCGCTACAGGGCGCTCCCATTCGCCATTTCAGGCTGCG	480
DB	3217	CACACCGCGCGCTTAATCGCGCTACAGGGCGCTCCCATTCGCCATTTCAGGCTGCG	3158
QY	481	CAACTGTTGGGAAGGCGATCGGTGGGCGCTCTTCGCTATTACGCCAGCTGGCGAAAGG	540
DB	3157	CAACTGTTGGGAAGGCGATCGGTGGGCGCTCTTCGCTATTACGCCAGCTGGCGAAAGG	3098
QY	541	GGGATGTCGCAAGCGGATTAAGTTGGGTAAAGCGGAGTTCCTCCAGTCAAGCTGTTG	600
DB	3097	GGGATGTCGCAAGCGGATTAAGTTGGGTAAAGCGGAGTTCCTCCAGTCAAGCTGTTG	3038
QY	601	TAAACGAGCGGCGAGTGGCGGCTCTTCATTACGTTTGTGAACCCGTCGAGGACGG	660
DB	3037	TAAACGAGCGGCGAGTGGCGGCTCTTCATTACGTTTGTGAACCCGTCGAGGACGG	2978
QY	661	GCAGACTCGCGGTGCAAAATGTGTTTACAGCGGTGATGGAGCAGATGAAGATGTCGACAC	720
DB	2977	CGCGGTGGCGG-----CGCTCTAGAACTAGTGGATCCCC--CGG	2939
QY	721	GCTGCAGACACCGACGCTAGATTAAACCTTAGAAAGATATCATATTGTGACGTACGTTAA	780
DB	2938	GCTGCAGGAATCTTATTGTTGATGTTTCATCCATGCCATGTGTAAATCCCGCAGCAGTTA	2879
QY	781	AGATAATCATGCTAAATTTGACGATGGGATCTGTAAATACGACTCCTATATAGGCGAAT	840
DB	2878	CAAACTCAAGAGGACCATGTGGTCAACGCTTTTCGTTGGGATCTTTTCGAAAGGGCAGATT	2819

QY	841	TGGGTACCGGGCCCCCTCGAGGTGCGAGCGGTATCGATAAGCTTGATATCGAATTCCTGC	900
DB	2818	GTGTCGACAGGTAATGTGTTGCTGTGTAAGGACAGGGCCATCGCCAAATGGAGTATTTT	2759
QY	901	AGCCCCGGGGATCCACTAGTTCTAGAGCGGCGCCACCGCGGTGGAGCTCCAGCTTTTGT	960
DB	2758	GTTGATAATGTGCTGTAGTTGAAACGATCCATCTTCAATGTTGTGGCGAATTTGAAGT	2699
QY	961	TCCCTTTAGTGAGGGTTAATTAGATCCCATGGTCAATTTTATCGCAGACTATCTTTCTAG	1020
DB	2698	TAGCTTTGATTCCATTTCTTTTGTCTGCGTGATGTATACATTTGTGTAG-TTATAG	2640
QY	1021	GGTTAATCTAGCTGCATCAGGATCATCTGTCGGGTCTTTTTCGGGCTCAGTCATCGCC	1080
DB	2639	TTGTATCTCAGTTTGTGTCGGAATGTTTTCCTCTTTTAAATCAATACCTTTTAAAC	2580
QY	1081	CAAGCTGGCGCTATCTGGGCATCGGGAGGAAGAGCCGCTGCTTTTCCCGGAGGTG	1140
DB	2579	TGATACGATTAAACAGGGTATCACCTTCAAACTTGACTTCAGCAGCGCTCTGTAGTTC	2520
QY	1141	AAGCGGCATGGAAGAGTTTGCAGGATGACTGTCTGCAATTGACGTTGAGCGGAAAC	1200
DB	2519	CGTCACTTTGAAAGATATAGTGTCTCTGTACATAACCTTCGGGCATGGCACTCTTG	2460
QY	1201	GCAGTTTACCATGATGTTCCGGGAAGGTGGCCATGCAAGCTTTTAACGGTGAACCTGT	1260
DB	2459	AAAAAGTCAATCCGCTTTCATATGATCCGATTAACGGGAAAGCATTTGAACCATAGAG	2400
QY	1261	TCGTTTCAGGCCACCTGGGATACAGTTCTGTCGGCTTTTTCGGACACAGTTCCTGGATGG	1320
DB	2399	AAAGTAGTGACAAAGTTTGGCCATGGAAACAGGTAGTTTTCAGTAGTGCAAAATTA	2340
QY	1321	TCAGCCGGAAGCGCATCAGCAACCCGAAACAATAACGGCGACAGCCGGAACCTGCGTSCCG	1380
DB	2339	AGGGTAAGCTTTTCGTAATGTAGCATCACCTTCACCTCTCCACTGACAGAAATTTGTGC	2280
QY	1381	GTGTGAGATTAAATGACAGCGGTGGCGCTGGGATATTAATGTCAGCAGGAGCGGTATC	1440
DB	2279	CCATTAACATCACTCACTTAATTTCAACAGAAATTTGGGACAACTCCAGTGAAGTTCCTCT	2220
QY	1441	CTGGCTGGATCGCGAGAAATGAGATGATACCCGCTGAGTTACCCGGCGGCGCGCTT	1500
DB	2219	C---CTTTACTCATCGGTACCCAGCTTTTGTTCCTTTTGTAGTGGGTTAATTCGCGCTT	2163
QY	1501	GGCGTAATCATGCTCATAGCTGTTCTGTGTAATTTGTTATCCGCTCACAAATTCACA	1560
DB	2162	GGCGTAATCATGCTCATAGCTGTTCTGTGTAATTTGTTATCCGCTCACAAATTCACA	2103
QY	1561	CAACATACGAGCCGGAAGCATAAAGTGTAAAGCTGGGTGCTTAATGATGAGCTTAAC	1620
DB	2102	CAACATACGAGCCGGAAGCATAAAGTGTAAAGCTGGGTGCTTAATGATGAGCTTAAC	2043
QY	1621	CACATTAATGCTGTCGCTCATCTGCGCTTTCAGTCCGGGAAACCTTCGTCGCGAGCT	1680
DB	2042	CACATTAATGCTGTCGCTCATCTGCGCTTTCAGTCCGGGAAACCTTCGTCGCGAGCT	1983
QY	1681	GCATTAATGAATCGGCAACCGCGGAGAGCGGTGTTGCGTATTTGGGCGCTCTTCCGC	1740
DB	1982	GCATTAATGAATCGGCAACCGCGGAGAGCGGTGTTGCGTATTTGGGCGCTCTTCCGC	1923
QY	1741	TTCTCGCTCATGCTGCTCGCTCGGTGCTGTTGGCTGCGGCGAGCGGTATCAGCTCA	1800
DB	1922	TTCTCGCTCATGCTGCTCGGTGCTGTTGGCTGCGGCGAGCGGTATCAGCTCA	1863
QY	1801	CTCAAGCGGTAAATACGTTATCACAGAAATCAGGGTAACCGAGGAAAGCAATG	1860
DB	1862	CTCAAGCGGTAAATACGTTATCACAGAAATCAGGGTAACCGAGGAAAGCAATG	1803
QY	1861	AGCAAAAGCGCAGCAAAAGCCAGGAAACCGTAAAGAGCGCGTTCCTGCGCTTTTCCA	1920
DB	1802	AGCAAAAGCGCAGCAAAAGCCAGGAAACCGTAAAGAGCGCGTTCCTGCGCTTTTCCA	1743

QY 1921 TAGGCTCCGCCCCCTGACGAGCATCACAAAATCGAGCTCAAGTCAGAGGTGGCGAAA 1980
Db 1742 TAGGCTCCGCCCCCTGACGAGCATCACAAAATCGAGCTCAAGTCAGAGGTGGCGAAA 1683
QY 1981 CCCGACAGACTATAAAGATACACAGCGCTTTCCCTCTGGAAGCTCCCTCGTGGCTCTCC 2040
Db 1682 CCCGACAGACTATAAAGATACACAGCGCTTTCCCTCTGGAAGCTCCCTCGTGGCTCTCC 1623
QY 2041 TGTTCCGACCCCTGCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCCGGAAGCGTGGC 2100
Db 1622 TGTTCCGACCCCTGCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCCGGAAGCGTGGC 1563
QY 2101 GCTTTCTCATAGCTCACGCTGPAAGTATCTCAGTTTCGGTGTAGGTGCTTCGCTCAAGCT 2160
Db 1562 GCTTTCTCATAGCTCACGCTGPAAGTATCTCAGTTTCGGTGTAGGTGCTTCGCTCAAGCT 1503
QY 2161 GGGCTGTGACGACGAACCCCGCTTACGCCGACCGCTGCGCTTATCCGTTAACTATCG 2220
Db 1502 GGGCTGTGACGACGAACCCCGCTTACGCCGACCGCTGCGCTTATCCGTTAACTATCG 1443
QY 2221 TCTTGAGTCCAAACCCGGTAAGACACGACTTATCGGCACCTGGCAGCAGCCACTGGTAACAG 2280
Db 1442 TCTTGAGTCCAAACCCGGTAAGACACGACTTATCGGCACCTGGCAGCAGCCACTGGTAACAG 1383
QY 2281 GATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTTAACTA 2340
Db 1382 GATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTTAACTA 1323
QY 2341 CGGCTACACTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGG 2400
Db 1322 CGGCTACACTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGG 1263
QY 2401 AAAAAGAGTTGTTAGTCTTGATCCGCAAAACAAACACACCGCTGTAGCGGTGGTTTTT 2460
Db 1262 AAAAAGAGTTGTTAGTCTTGATCCGCAAAACAAACACACCGCTGTAGCGGTGGTTTTT 1203
QY 2461 TGTTTGAAGCAGCAGATTAACGCGAGAAAAAAGGATCTCAAGAAGATCTTTTGATCTT 2520
Db 1143 TGTTTGAAGCAGCAGATTAACGCGAGAAAAAAGGATCTCAAGAAGATCTTTTGATCTT 1143
QY 2521 TTCTACGGGGTCTGACGCTCAGTGAAGGAAACCTCAGGTTAAGGGATTTTGGTCATGAG 2580
Db 1142 TTCTACGGGGTCTGACGCTCAGTGAAGGAAACCTCAGGTTAAGGGATTTTGGTCATGAG 1083
QY 2581 ATTATCAAAAAGGATCTTCACTAGATCTTTAAATTAATAAATGAAGTTTAAATCAAT 2640
Db 1082 ATTATCAAAAAGGATCTTCACTAGATCTTTAAATTAATAAATGAAGTTTAAATCAAT 1023
QY 2641 CTAAGTATATATAGTAAACTTGGTCTGACAGTTTACCAATGCTTAAATCAGTGGGCACC 2700
Db 1022 CTAAGTATATATAGTAAACTTGGTCTGACAGTTTACCAATGCTTAAATCAGTGGGCACC 963
QY 2701 TATCTCAGGATCTGTCTATTCTTGGTTTCATCCNATGTTGCTGACTCCCGTCTGTAGAT 2760
Db 962 TATCTCAGGATCTGTCTATTCTTGGTTTCATCCNATGTTGCTGACTCCCGTCTGTAGAT 903
QY 2761 AACTACGATACGGGAGGCTTACCATCTGGCCCCAGTGTGCAATGATACCCGAGACCC 2820
Db 902 AACTACGATACGGGAGGCTTACCATCTGGCCCCAGTGTGCAATGATACCCGAGACCC 843
QY 2821 ACGCTCACCGGCTCCAGATTTTATCAGCAATAAACCAGCCAGCCGGAAGGGCGGAGCGAG 2880
Db 842 ACGCTCACCGGCTCCAGATTTTATCAGCAATAAACCAGCCAGCCGGAAGGGCGGAGCGAG 783
QY 2881 AAGTGGTCTGCAACTTTATCGGCTCCATCCAGTCTATTAAATTTGTCGGGAAGCTAG 2940
Db 782 AAGTGGTCTGCAACTTTATCGGCTCCATCCAGTCTATTAAATTTGTCGGGAAGCTAG 723
QY 2941 AGTAAGTAGTTCCGCGAGTTAATAGTTTTCGCAACGTTGTGCAATTTGCTACAGGCATCGT 3000
Db 722 AGTAAGTAGTTCCGCGAGTTAATAGTTTTCGCAACGTTGTGCAATTTGCTACAGGCATCGT 663
QY 3001 GGTGTACGCTCGTCTGTTGGTATGGCTTCAATTCAGCTCCGTTCCCAACGATCAAGGCG 3060

Db 662 GGTGTACGCTCGTCTGTTGGTATGGCTTCAATTCAGCTCCGTTCCCAACGATCAAGGCG 603
QY 3061 AGTTACATGATCCCCCATGTTGTGCAAAAGAGCGGTAGCTCCTTCGGTCTCCGATCGT 3120
Db 602 AGTTACATGATCCCCCATGTTGTGCAAAAGAGCGGTAGCTCCTTCGGTCTCCGATCGT 543
QY 3121 TGTCAAGAGTAAGTTGGCGCAGTGTATCACTCATGTGTATGGCAGCACTGCATAATTC 3180
Db 542 TGTCAAGAGTAAGTTGGCGCAGTGTATCACTCATGTGTATGGCAGCACTGCATAATTC 483
QY 3181 TCTTACTGTCTATGCAATCCGTAAGATGCTTTTCTGTGATCGTGTAGTACTCAACCAAGTC 3240
Db 482 TCTTACTGTCTATGCAATCCGTAAGATGCTTTTCTGTGATCGTGTAGTACTCAACCAAGTC 423
QY 3241 ATTCTGAGAATAGTGTATGCGCGACCGAGTTGCTTTCGCCGCGTCAATACGGGATAA 3300
Db 422 ATTCTGAGAATAGTGTATGCGCGACCGAGTTGCTTTCGCCGCGTCAATACGGGATAA 363
QY 3301 TACCGCGCACATAGCAGAACTTTAAAAAGTGTCTCATTTGGAAAAACGTTCTTCGGGCG 3360
Db 362 TACCGCGCACATAGCAGAACTTTAAAAAGTGTCTCATTTGGAAAAACGTTCTTCGGGCG 303
QY 3361 AAACTCTCAAGGATCTTACCGCTGTGTGAGATCCAGTTGATGTAACCACTCGTGCACC 3420
Db 302 AAACTCTCAAGGATCTTACCGCTGTGTGAGATCCAGTTGATGTAACCACTCGTGCACC 243
QY 3421 CAACCTGATCTTCAGCATCTTTTACTTTCACCAAGCGTTTCTGGTGAGCAAAAACAGGAAG 3480
Db 242 CAACCTGATCTTCAGCATCTTTTACTTTCACCAAGCGTTTCTGGTGAGCAAAAACAGGAAG 183
QY 3481 GCAAAATCCGCAAAAAGGAAATAAGGCGCAGACGGAATAATGTTGAATCTCATACTCTT 3540
Db 182 GCAAAATCCGCAAAAAGGAAATAAGGCGCAGACGGAATAATGTTGAATCTCATACTCTT 123
QY 3541 CCTTTTCAATATTATTGAAGCAATTTATCAGGGTTATTGTCTCATGAGGGGATACATATT 3600
Db 122 CCTTTTCAATATTATTGAAGCAATTTATCAGGGTTATTGTCTCATGAGGGGATACATATT 63
QY 3601 TGAATGTATTTAGAAAAATAAACAATAGGGGTTCCGCGCACATTTCCCGAAAAAGTGCC 3660
Db 62 TGAATGTATTTAGAAAAATAAACAATAGGGGTTCCGCGCACATTTCCCGAAAAAGTGCC 3

RESULT 12

US-10-280-913A-17/c
; Sequence 17, Application US/10280913A
; Publication No. US20040110130A1
; GENERAL INFORMATION:
; APPLICANT: Large Scale Biology Corporation
; TITLE OF INVENTION: POPULATION OF POLYNUCLEOTIDE SEQUENCE VARIANTS
; FILE REFERENCE: LSBC-PLG5392-CIP
; CURRENT APPLICATION NUMBER: US/10/280,913A
; CURRENT FILING DATE: 2002-10-25
; PRIOR APPLICATION NUMBER: 60/402,342
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 10/066,390
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 60/268,785
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/266,386
; PRIOR FILING DATE: 2002-02-02
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 3637
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:

OTHER INFORMATION: Encodes cycle 3 GFP
US-10-280-913A-17

Query Match 72.8%; Score 2664.8; DB 18; Length 3637;
Best Local Similarity 83.8%; Pred. No. 0;
Matches 3070; Conservative 0; Mismatches 567; Indels 25; Gaps 4;
QY 1 CTAATTTGTAAGGCTTAATATTTTGTGTTAAATTCGGGTTAAATTTTGTGTTAAATTCAGCTC 60
DB 3637 CTAATTTGTAAGGCTTAATATTTTGTGTTAAATTCGGGTTAAATTTTGTGTTAAATTCAGCTC 3578
QY 61 ATTTTAAACCAATAGCGCGAAATCGGCAAAATCCCTTATAATCAAAAGAAATAGACCGA 120
DB 3577 ATTTTAAACCAATAGCGCGAAATCGGCAAAATCCCTTATAATCAAAAGAAATAGACCGA 3518
QY 121 GATAGGTTGAGTGTGTTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGAAAGTGGACTC 180
DB 3517 GATAGGTTGAGTGTGTTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGAAAGTGGACTC 3458
QY 181 CAAAGTCAAAAGGCGGAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCAC 240
DB 3457 CAAAGTCAAAAGGCGGAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCAC 3398
QY 241 CTAATCAAGTCTTTTGGGTCGAGTCCGTAAGCACTAAATCGGAACCTTAAAGGGAG 300
DB 3397 CTAATCAAGTCTTTTGGGTCGAGTCCGTAAGCACTAAATCGGAACCTTAAAGGGAG 3338
QY 301 CCCCCGATTTAGAGCTTTGACGGGAAAGCCGCGAACGTTGGCGAGAAAGAAAGGAAAGAA 360
DB 3337 CCCCCGATTTAGAGCTTTGACGGGAAAGCCGCGAACGTTGGCGAGAAAGGAAAGGAAAGAA 3278
QY 361 AGCGAAAGAGCGGGCGCTAGGGCGTGGCAAGTGTAGCGGTCACTCGCGGTAAACCA 420
DB 3277 AGCGAAAGAGCGGGCGCTAGGGCGTGGCAAGTGTAGCGGTCACTCGCGGTAAACCA 3218
QY 421 CACACCGCGCGCTTAATGCGCGCTACAGGCGGCTCCCATTCGCCATTCAGGCTGCG 480
DB 3217 CACACCGCGCGCTTAATGCGCGCTACAGGCGGCTCCCATTCGCCATTCAGGCTGCG 3158
QY 481 CAACTGTTCGGGAAGGCGATCGGTGGCGCTCTTCGCTATTACGCCAGCTGGCGAAAGG 540
DB 3157 CAACTGTTCGGGAAGGCGATCGGTGGCGCTCTTCGCTATTACGCCAGCTGGCGAAAGG 3098
QY 541 GGGATGTGTCGAAGCGAATTAAGTTGGGTAAAGCGCAGGCTTTCCAGTCAAGCGTTG 600
DB 3097 GGGATGTGTCGAAGCGAATTAAGTTGGGTAAAGCGCAGGCTTTCCAGTCAAGCGTTG 3038
QY 601 TAAACGACGCGCAGTACGCGGCTCGTTTCACTACGTTTGTGAACCCGCTGGAGACGG 660
DB 3037 TAAACGACGCGCAGTACGCGGCTCGTTTCACTACGTTTGTGAACCCGCTGGAGTCCA 2978
QY 661 GCAGACTCGCGTGCAAAATGTGTTTACAGCGTGTAGGAGCAGATGAAGATGCTCGACAC 720
DB 2977 CGCGGTGCGG-----CGGCTCTGAACCTAGTGATCCCC--CGG 2939
QY 721 GCTGCAGAAACCGCAGCTAGATTAAACCTAGAAAGATAATCATATTGTGACCGTAA 780
DB 2938 GCTGCAGGAAATCTTATTGTTATAGTTTCATCCATGCCATGTGTAATCCCGCAGCAGTTA 2879
QY 781 AGATAATCATGCGTAAATTTGACGATGGATCTGTAATACGACTCACTATAGGGCGAAT 840
DB 2878 CAAACTCAAGAAGGACCATGTGCTCAAGCTTTTCGTTGGGATCTTTTCGAAAGGGCAGATT 2819
QY 841 TGGGTACCGGGCCCCCTCGAGTTCGAGGTATCGATAAGCTTGATATCGAATTCCTGC 900
DB 2818 GTGTCACAGGTAATGGTTGTCTGGTAAAGGACAGGGCCATCGCCAAATTGGAGTATTTT 2759
QY 901 AGCCCGGGGATCCACTAGTTCTAGAGCGCGCCGACACCGCGGTGGAGCTCCAGCTTTTGT 960
DB 2758 GTTGATAATGGTCTGCTAGTTGAAACGATCCATCTTCAATGTTGTGGCGAATTTGAAGT 2699
QY 961 TCCCTTTAGTGGGTTAATTAGATCCCATGCGTCAATTTTACGACAGACTATCTTTCTAG 1020
DB 1020 TCCCTTTAGTGGGTTAATTAGATCCCATGCGTCAATTTTACGACAGACTATCTTTCTAG 1563

DB 2698 TAGCTTTGATCCATTCCTTTTGTGTCGCGTATGTATACATTTGTGTGAG-TTATAG 2640
QY 1021 GGTAACTAGCTGCAATCAGGATCATATCGTCGGGCTTTTTCGCGCTCAGTCATGCGC 1080
DB 2639 TTGTACTCGAGTTTGTGTCGAGAAATGTTTCCATCTTCTTTAAATAACAATACCTTTTAAAC 2580
QY 1081 CAAAGTGGCGCTATCTGGCATCGGGAGGAAGACCGTCCCTTTTCCCGGAGGTG 1140
DB 2579 TCGATACGATTAACAAGGGTATCACTTCAAACTTGCACGCGCTTTGTGTAGTTT 2520
QY 1141 AAGCGCATGGAAGAGTTTGCAGGATGACTGCTGTGCAATGACGTTGAGCGAAAC 1200
DB 2519 CGGTATCTTTGAAAGATATAGTCTCTGTACATAAACCCTTCGGCATGGCACTCTTG 2460
QY 1201 GCACGTTTACATGATGATTCGGGAAGGTGGCCATGCAACGCTTTTAAACGTTGAACGTGT 1260
DB 2459 AAAAGTCTATCGCTTTTCAATATGATCCGATTAACGGGAAAGCATTTGAACACCATTAAGAG 2400
QY 1261 TCGTTCAGGCACTGGGATACAGTTCGTCGGGCTTTTCCGGACACAGTTCGGATGG 1320
DB 2399 AAAGTAGTGACAAGTGTGGCCATGGAAACAGGTAGTTTCCAGTAGTGCAAAATAAATTTA 2340
QY 1321 TCAGCCCGAAGCGCATCAGCAACCCGAAACAATAACCGCGACAGCCGGAACCTGCCGTGCCG 1380
DB 2339 AGGTAAAGCTTTCCGTATGTAGCATCACCTTCACTCTCCACTGACAGAAATTTGTGC 2280
QY 1381 GTGTGAGATTAATGACAGCGGTGGCGCTGGGATTAATACGTACGCGAGGACGGGTATC 1440
DB 2279 CCATTAACATCACCATCTAATTTCAACAAGAAATTTGGGACAACCTCCAGTGAAAGTCTTCT 2220
QY 1441 CTGGCTGGATGGCGCAAAATGGATACCGGTGAGTTACCGGGGGCGGCGCTT 1500
DB 2219 C---CTTACTCATCGGTACCCAGCTTTTGTTCCTTTAGTGAGGTTAATTTGCGGCTT 2163
QY 1501 GCGGTAATCATGTCTATGCTGTTCTGTGTGAAATGTTATCCGCTCACAAATTCACA 1560
DB 2162 GCGGTAATCATGTCTATGCTGTTCTGTGTGAAATGTTATCCGCTCACAAATTCACA 2103
QY 1561 CAACTACAGAGCGGAAAGCATAAAGTGTAAAGCTGGGGTGCCTAAATGATGAGCTAACT 1620
DB 2102 CAACTACAGAGCGGAAAGCATAAAGTGTAAAGCTGGGGTGCCTAAATGATGAGCTAACT 2043
QY 1621 CACTTAATTTGGTTGCGCTCACTGCGCTTTCAGTTCGGGAAACCTGTCTGTCAGCT 1680
DB 2042 CACTTAATTTGGTTGCGCTCACTGCGCTTTCAGTTCGGGAAACCTGTCTGTCAGCT 1983
QY 1681 GCATTAATGAATCGGCAACGCGCGGAGAGCGGTTTTCGTTATTTGGCGCTTCTCCGC 1740
DB 1982 GCATTAATGAATCGGCAACGCGCGGAGAGCGGTTTTCGTTATTTGGCGCTTCTCCGC 1923
QY 1741 TTCTCTGCTCACTGACTCGCTCGCTCGCTTTCGCTTCGGCGAGCGGTATCAGCTCA 1800
DB 1922 TTCTCTGCTCACTGACTCGCTCGCTCGCTTTCGCTTCGGCGAGCGGTATCAGCTCA 1863
QY 1801 CTCAAAGGCGGTAAATACGTTATTCAGAAATCAGGGGATAACGAGGAAAGAAATGTC 1860
DB 1862 CTCAAAGGCGGTAAATACGTTATTCAGAAATCAGGGGATAACGAGGAAAGAAATGTC 1803
QY 1861 AGCAAAAGCCAGCAAAAGCCAGGAACCGTAAAGGCGCGGTTCTGCTGCGTTTTCCTCA 1920
DB 1802 AGCAAAAGCCAGCAAAAGCCAGGAACCGTAAAGGCGCGGTTCTGCTGCGTTTTCCTCA 1743
QY 1921 TAGGCTCCCGCCCTTCAGCAGCATCAAAAATCGACGCTCAAGTCAAGAGTGGGGA 1980
DB 1742 TAGGCTCCCGCCCTTCAGCAGCATCAAAAATCGACGCTCAAGTCAAGAGTGGGGA 1683
QY 1981 CCCGACAGGACTATAAAGATAACGCGTTTCCCTCTGGAAGCTCCCTGTCGCTCTCC 2040
DB 1682 CCCGACAGGACTATAAAGATAACGCGTTTCCCTCTGGAAGCTCCCTGTCGCTCTCC 1623
QY 2041 TGTTCGACCCCTGCGCTTACCGGATACCTGTGCGCTTTCTCCCTTCGGGAGCGTGGC 2100
DB 1622 TGTTCGACCCCTGCGCTTACCGGATACCTGTGCGCTTTCTCCCTTCGGGAGCGTGGC 1563

```
QY 2101 GCTTTCTCATAGCTCACGCTGATAGTATCTCAGTTCCGTTAGTTCGTTCCGCTCAAGCT 2160
DB |||||||
1562 GCTTTCTCATAGCTCACGCTGATAGTATCTCAGTTCCGTTAGTTCGTTCCGCTCAAGCT 1503
QY 2161 GGGCTGTGTGCAAGAACCCCGTTCAGCCCGACCGCTGCGCTTATCCGGTAACTATCG 2220
DB |||||||
1502 GGGCTGTGTGCAAGAACCCCGTTCAGCCCGACCGCTGCGCTTATCCGGTAACTATCG 1443
QY 2221 TCTTCAGTCCACCCGCTAAGACACGACTTATCGCACTGGCAGCAGCCACTGCTAAACAG 2280
DB |||||||
1442 TCTTCAGTCCACCCGCTAAGACACGACTTATCGCACTGGCAGCAGCCACTGCTAAACAG 1383
QY 2281 GATTAGCAGACGAGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGCCTAACTA 2340
DB |||||||
1382 GATTAGCAGACGAGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGCCTAACTA 1323
QY 2341 CGGCTACACTAGAAGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTCCTTCGG 2400
DB |||||||
1322 CGGCTACACTAGAAGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTCCTTCGG 1263
QY 2401 AAAAGAGTTGGTAGCTCTTGATCCGGCAAAACCAACCAACCGCTGCTAGGGTGGTTTTT 2460
DB |||||||
1262 AAAAGAGTTGGTAGCTCTTGATCCGGCAAAACCAACCAACCGCTGCTAGGGTGGTTTTT 1203
QY 2461 TGTTCGAAAGCAGAGATTACCGCAGAAAGAAAGGATCTCAAGAGATCTCTTTGATCTT 2520
DB |||||||
1202 TGTTCGAAAGCAGAGATTACCGCAGAAAGAAAGGATCTCAAGAGATCTCTTTGATCTT 1143
QY 2521 TTCTACGGGGTCTGACGCTCAGTGGAAAGAAACTCAGCTTAAAGGATTTTGGTCAATGAG 2580
DB |||||||
1142 TTCTACGGGGTCTGACGCTCAGTGGAAAGAAACTCAGCTTAAAGGATTTTGGTCAATGAG 1083
QY 2581 ATTATCAAAAGGATCTTCACTAGATCTTTTAAATTAATAATCAAGTTTAAATCAAT 2640
DB |||||||
1082 ATTATCAAAAGGATCTTCACTAGATCTTTTAAATTAATAATCAAGTTTAAATCAAT 1023
QY 2641 CTAAGATATATAGTAAACTTTGCTGACAGTTACCAATGCTTTAATCAGTGAGGCACC 2700
DB |||||||
1022 CTAAGATATATAGTAAACTTTGCTGACAGTTACCAATGCTTTAATCAGTGAGGCACC 963
QY 2701 TATCTCAGCGATCTGCTATTTGTTTCATPCATAGTTGCTGACTCCCGCGTGTAGAT 2760
DB |||||||
962 TATCTCAGCGATCTGCTATTTGTTTCATPCATAGTTGCTGACTCCCGCGTGTAGAT 903
QY 2761 AACTACGATACGGGAGGCTTACCATCTGGCCAGTCTGCAATGATACCCGAGACCC 2820
DB |||||||
902 AACTACGATACGGGAGGCTTACCATCTGGCCAGTCTGCAATGATACCCGAGACCC 843
QY 2821 ACGCTCACCGGCTCCAGATTTATCAGCAATAAACCCAGCAGCCGGAAGGCGCGAGCAG 2880
DB |||||||
842 ACGCTCACCGGCTCCAGATTTATCAGCAATAAACCCAGCAGCCGGAAGGCGCGAGCAG 783
QY 2881 AAGTGGTCTCGCACTTTATCCGCTCCATCCAGTCTATTAAATGTTCCCGGGAAGCTAG 2940
DB |||||||
782 AAGTGGTCTCGCACTTTATCCGCTCCATCCAGTCTATTAAATGTTCCCGGGAAGCTAG 723
QY 2941 AGTAAGTAGTTCGCCAGTTAATAGTTTGGCAAGCTTGTGCCATTGCTACAGGCATCGT 3000
DB |||||||
722 AGTAAGTAGTTCGCCAGTTAATAGTTTGGCAAGCTTGTGCCATTGCTACAGGCATCGT 663
QY 3001 GGTGTACGCTCGTCTGTTGGTATGGCTTCAATCAGCTCCGTTCCCAACCATCAAGGCG 3060
DB |||||||
662 GGTGTACGCTCGTCTGTTGGTATGGCTTCAATCAGCTCCGTTCCCAACCATCAAGGCG 603
QY 3061 AGTTACATGATCCCCCATGTTGTGCAAAAAGCGGTTAGTCTCTTCGTTCTCCGATCGT 3120
DB |||||||
602 AGTTACATGATCCCCCATGTTGTGCAAAAAGCGGTTAGTCTCTTCGTTCTCCGATCGT 543
QY 3121 TGTCAAGATAGTTGGCGGAGTGTATCACTCATGTTATGCGCAGCAGCTGCATAATTC 3180
DB |||||||
542 TGTCAAGATAGTTGGCGGAGTGTATCACTCATGTTATGGCAGCAGCTGCATAATTC 483
```

```
QY 3181 TCTTACTGTCTGCGCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTC 3240
DB |||||||
482 TCTTACTGTCTGCGCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTC 423
QY 3241 ATTCTGAGAATAGTGTATGCGCGGACCGAGTTGCTCTTGGCCGCGGTCAATACGGGATAA 3300
DB |||||||
422 ATTCTGAGAATAGTGTATGCGCGGACCGAGTTGCTCTTGGCCGCGGTCAATACGGGATAA 363
QY 3301 TACCGCGCACATAGCAGAACTTTAAAGTGTCTCATCATTTGGAAAACTTCTTCGGGGCG 3360
DB |||||||
362 TACCGCGCACATAGCAGAACTTTAAAGTGTCTCATCATTTGGAAAACTTCTTCGGGGCG 303
QY 3361 AAAAATCTCAAGGATCTTTACCGCTGTTGAGATCCAGTTTCGATGTAAACCCACTCGTGCACC 3420
DB |||||||
302 AAAAATCTCAAGGATCTTTACCGCTGTTGAGATCCAGTTTCGATGTAAACCCACTCGTGCACC 243
QY 3421 CAACTGATCTTCAGCATCTTTTCACTTTCCAGCGCTTTCTGGGTGAGCAAAAACAGGAAG 3480
DB |||||||
242 CAACTGATCTTCAGCATCTTTTCACTTTCCAGCGCTTTCTGGGTGAGCAAAAACAGGAAG 183
QY 3481 GCAAAATCCGCAAAAAGGGAATAAGGCGCACGGAATAATGTTGAATACTCATACTCTT 3540
DB |||||||
182 GAAAATCCGCAAAAAGGGAATAAGGCGCACGGAATAATGTTGAATACTCATACTCTT 123
QY 3541 CCTTTTCAATATTTATTTGAAGCATTTTATCAGGGTTATTGTCTCATGAGCGGATACATATT 3600
DB |||||||
122 CCTTTTCAATATTTATTTGAAGCATTTTATCAGGGTTATTGTCTCATGAGCGGATACATATT 63
QY 3601 TGAATGTAATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCGGAAAAAGTGCC 3660
DB |||||||
62 TGAATGTAATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCGGAAAAAGTGCC 3
QY 3661 AC 3662
DB 2 AC 1
```

RESULT 13
US-10-684-134-17/c
; Sequence 17, Application US/10684134
; Publication No. US20040142433A1
; GENERAL INFORMATION:
; APPLICANT: Large Scale Biology Corporation
; TITLE OF INVENTION: POPULATION OF POLYNUCLEOTIDE SEQUENCE VARIANTS
; FILE REFERENCE: LSBC-PLG5392-CIP
; CURRENT APPLICATION NUMBER: US/10/684,134
; CURRENT FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: 60/402,342
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 10/066,390
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 60/268,785
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/266,386
; PRIOR FILING DATE: 2002-02-02
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 3637
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Encodes cycle 3 GFP
US-10-684-134-17

Query Match 72.8%; Score 2664.8; DB 18; Length 3637;
Best Local Similarity 83.8%; Pred. No. 0;
Matches 3070; Conservative 0; Mismatches 567; Indels 25; Gaps 4;
QY 1 CTAATTTCTAGCGTTAATATTTTCTTAAATTCGCGTTAAATTTTCTTAAATCAGCTC 60
DB 3637 CTAATTTCTAGCGTTAATATTTTCTTAAATTCGCGTTAAATTTTCTTAAATCAGCTC 3578

QY 61 ATTTTAAACCAATAGCGCGAATCGGCAAAATCCCTTATAATCAAAAGATAGACCGA 120
DB 3577 ATTTTAAACCAATAGCGCGAATCGGCAAAATCCCTTATAATCAAAAGATAGACCGA 3518
QY 121 GATAGGGTTGAGTGTGTTCCAGTTTGGAAACAGAGTCCACTATTAAAGAAAGTGGACTC 180
DB 3517 GATAGGGTTGAGTGTGTTCCAGTTTGGAAACAGAGTCCACTATTAAAGAAAGTGGACTC 3458
QY 181 CAAAGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCAATCACC 240
DB 3457 CAAAGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCAATCACC 3398
QY 241 CTAATCAAGTTTTTGGGTGAGGTGCGGTAAAGCACTAAATCGGAAACCTTAAGGGAG 300
DB 3397 CTAATCAAGTTTTTGGGTGAGGTGCGGTAAAGCACTAAATCGGAAACCTTAAGGGAG 3338
QY 301 CCCCCGATTTAGAGCTTTCAGGGGAAGCGGCAACGTTGGCGAGAAAGGAAGGAAGA 360
DB 3337 CCCCCGATTTAGAGCTTTCAGGGGAAGCGGCAACGTTGGCGAGAAAGGAAGGAAGA 3278
QY 361 AGCGAAAGAGCGGGCGCTTAGGGCGCTGGCAAGTGTAGCGGTCAACGCTCGCGCTAACCCAC 420
DB 3277 AGCGAAAGAGCGGGCGCTTAGGGCGCTGGCAAGTGTAGCGGTCAACGCTCGCGCTAACCCAC 3218
QY 421 CACACCGCGCGCTTAATGCGCGCTACAGGGCGGTCCTTCCGCTTACGCTGCG 480
DB 3217 CACACCGCGCGCTTAATGCGCGCTACAGGGCGGTCCTTCCGCTTACGCTGCG 3158
QY 481 CAACTGTTGGGAAGGGCGATCGGTGGGGCTCTTCGCTATTACGGCACTGCGGAAAGG 540
DB 3157 CAACTGTTGGGAAGGGCGATCGGTGGGGCTCTTCGCTATTACGGCACTGCGGAAAGG 3098
QY 541 GGGATGTGCTCAAGGGCGATTAAGTTGGGTAAAGCGCAGGGGTTTTCCAGTCAAGAGTTG 600
DB 3097 GGGATGTGCTCAAGGGCGATTAAGTTGGGTAAAGCGCAGGGGTTTTCCAGTCAAGAGTTG 3038
QY 601 TAAAGCAGCGGAGTGGCGGCTCGCTTCTTCACTGTTTGAACCGCTGGAGAGCG 660
DB 3037 TAAAGCAGCGGAGTGGCGGCTCGCTTCTTCACTGTTTGAACCGCTGGAGAGCG 2978
QY 661 GCAGACTCGCGTGCAAAATGTGTTTACAGCGTATGGAGCAGATGAAGATCTCGACAC 720
DB 2977 CCGGCTGCGG-----CGGCTCTAAGACTAGTGGATCCCC--CGG 2939
QY 721 GCTGCAGAACACGCGACTAGATTAAACCTTAGAAAGATAATCATATTGTGACGTACGTAA 780
DB 2938 GCTGCAGGAATCTTATTGTTGATGTTTCATCCATGCCATGTGTAATCCCGCAGCAGTTA 2879
QY 781 AGATAATCATGCGTAAATTTAGCGCATGGGATCTGTAATACGACTCACTATAGGGCGAAT 840
DB 2878 CAAACTCAAGAAAGGACCATGTGTCACGCTTTTCGTTGGGATCTTTTCGAAAGGGCAGATT 2819
QY 841 TGGGTACCGGGCCCCCTCGAGTTCGACGGTATCGATAAGCTTGATATCGAATTCCTGC 900
DB 2818 GTGTCGACAGGTAATGGTTGCTGTGTAAGGACAGGGCCATCGCCAAATGGAGATTATT 2759
QY 901 AGCCCGGGGGATCCACTAGTTCTTAGAGCGGCGGCCACCGCGGTGGAGCTCCAGCTTTTGT 960
DB 2758 GTTGATAATGGTCTGCTAGTTGAAACGATCCATCTTCAATGTTGTGGCGAATTTTGAAGT 2699
QY 961 TCCCTTTAGTAGGGTTAATAGATCCCATGGGTCAATTTTATAGCGAGACTATCTTTCTAG 1020
DB 2698 TAGCTTTGATTCATCTTTTGTGTTGCTGCGGTGATGATATCAATTTGTGTGAG-TTATAG 2640
QY 1021 GGTAACTTAGTGCATCAGGATCATATCGTGGGTCTTTTTCGGCTCAGTCACTCGCC 1080
DB 2639 TTGTACTCGAGTTTGTGTCGGAATGTTTTCATCTTCTTAAATCAATCTTTTAAAC 2580
QY 1081 CAAAGTGGCGCTATCTGGGCATCGGGGAGGAAGACCGGTGCTTTTCCCGCAGGGTTG 1140
DB 2579 TCGATACGATTAAACAGGGTATCACTTCAAACTTGACTTCAGCAGCGGCTTGTAGTTC 2520
QY 1141 AAGCGGCATGGAAAGATTTTGGCGAGGATGACTGCTGCTGCATGAGCTTGAAGCGAAGAC 1200

DB 2519 CCGTCATCTTTGAAAGATATAGTGGCTTCTGTATACATAACCTTCGGGCATAGCACTCTTG 2460
QY 1201 GCACGCTTTACCATGATGATTTCGGGAAGGTGTGGCCATGACACGCTTTTAAACGGTGAACGTGT 1260
DB 2459 AAAAAGTCATCGCTTTTCATATGATCCGGATAACGGGAAAGCAATGAACACCATTAAGAG 2400
QY 1261 TCGTTCAGGCCACCTGGGATACAGTTCGTGCGGCTTTTCGGGACACAGTTCGGGATGG 1320
DB 2399 AAAAGTAGTGACAAGTGTGGCCATGGAAACAGGTAGTGTTCCTCCAGTAGTGCAAAATAATTA 2340
QY 1321 TCAGCCCCGAAGCGCATCAGCAACCCGAAACAATACCGGCGACAGCCGGAACCTGCGGCGG 1380
DB 2339 AGGGTAAGCTTTCCGTATAGCATCACCTTCACCCCTCTCCACTGACAGAAATTTGTGC 2280
QY 1381 GTGTGCAGATTAAATGACAGCGGTGGCGCTGGGATATTACGTACGCGAGGACGGGTATC 1440
DB 2279 CCATTAACATACCATCTAATTTCAACAAATTTGGGACAACCTCCAGTGAAGTTCTTCT 2220
QY 1441 CTGGCTGGATGCGGAGAAATGGAATGGAATACCCCGTAGATTACCCGGGGGGCGGCTT 1500
DB 2219 C---CTTTACTCATCGGTACCCAGCTTTTGTTCCTTTAGTAGGGGTTAATTCGGGCTT 2163
QY 1501 GCGGTAATCATGCTCATAGCTGTTCTGTGTGAATTTGTTATCCGCTCACAAATTCACA 1560
DB 2162 GCGGTAATCATGCTCATAGCTGTTCTGTGTGAATTTGTTATCCGCTCACAAATTCACA 2103
QY 1561 CAACTATACAGCGCGGAAGCATAAAGTGTAAAGCTGGGGTGCCTAATGATGAGTGAAGT 1620
DB 2102 CAACTATACAGCGCGGAAGCATAAAGTGTAAAGCTGGGGTGCCTAATGATGAGTGAAGT 2043
QY 1621 CACTTAATGCTGCTCATCTGCTGCTCTCATGCTGCTTTCAGTCCGGAACCTGTGTCGAGCT 1680
DB 2042 CACTTAATGCTGCTGCTCTCATGCTGCTTTCAGTCCGGAACCTGTGTCGAGCT 1983
QY 1681 GCATTAATGAATCGGCAACCGCGGAGAGCGGTTTGCCTTATTTGGGGCTCTTCCGC 1740
DB 1982 GCATTAATGAATCGGCAACCGCGGAGAGCGGTTTGCCTTATTTGGGGCTCTTCCGC 1923
QY 1741 TTTCTCGCTCACTGACTCGCTCGCTCGCTCGTTCGGCTCGCGGAGCGGTATCAGCTCA 1800
DB 1922 TTTCTCGCTCACTGACTCGCTCGCTCGCTCGTTCGGCTCGCGGAGCGGTATCAGCTCA 1863
QY 1801 CTCAAAGCGGTAAATACGGTTATCAACAGATCAAGGATTAACGCGGAAAGAAATATGTG 1860
DB 1862 CTCAAAGCGGTAAATACGGTTATCAACAGATCAAGGATTAACGCGGAAAGAAATATGTG 1803
QY 1861 AGCAAAAGCCAGCAAAAGGCGCAGGAAACGTTAAAGGCGCGTTGCTGCGGTTTTCCTCA 1920
DB 1802 AGCAAAAGCCAGCAAAAGGCGCAGGAAACGTTAAAGGCGCGTTGCTGCGGTTTTCCTCA 1743
QY 1921 TAGGCTTCGCCCCCTCGAGCAGCATCAACAAATCGACGCTCAAGTCAAGGTGGCGGAAA 1980
DB 1742 TAGGCTTCGCCCCCTCGAGCAGCATCAACAAATCGACGCTCAAGTCAAGGTGGCGGAAA 1683
QY 1981 CCCGACAGGACTATAAAGATACAGCGGTTTCCCTCTGAAAGCTCCCTCGTGGCTCTCC 2040
DB 1682 CCCGACAGGACTATAAAGATACAGCGGTTTCCCTCTGAAAGCTCCCTCGTGGCTCTCC 1623
QY 2041 TGTTCCGACCCCTGCGCTTACCGGATACCTGTCCGCTTTCCTCCCTCGGAGCGGTGC 2100
DB 1622 TGTTCCGACCCCTGCGCTTACCGGATACCTGTCCGCTTTCCTCCCTCGGAGCGGTGC 1563
QY 2101 GCTTTCTCATAGCTCACTGCTGATAGGTATCTCAGTTCCGTTGATAGGTGCTTCGCTCAAGCT 2160
DB 1562 GCTTTCTCATAGCTCACTGCTGATAGGTATCTCAGTTCCGTTGATAGGTGCTTCGCTCAAGCT 1503
QY 2161 GGGCTGTGTGACGAAACCCCGCTTACGCCGAGCGCTCGGCTTATCCGGTAACTATCG 2220
DB 1502 GGGCTGTGTGACGAAACCCCGCTTACGCCGAGCGCTCGGCTTATCCGGTAACTATCG 1443
QY 2221 TCTTGAGTCCAAACCGGTTAAGACAGACTTATCGGCACTTGGCAGCGACCTGTTGAACAG 2280

Db 1442 TCTTGAGTCAACCCGGTAAGA CACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAG 1383
Qy 2281 GATTAGCAGAGAGGTATGTAGGCGGTCTACAGATCTTGAAGTGGTGGCCCTAACTA 2340
Db |||||
Db 1382 GATTAGCAGAGAGGTATGTAGGCGGTCTACAGATCTTGAAGTGGTGGCCCTAACTA 1323
Qy 2341 CGGCTACACTAGAGGACAGATATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGG 2400
Db |||||
Db 1322 CGGCTACACTAGAGGACAGATATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGG 1263
Qy 2401 AAAAGAGTTGGTACTCTTGATTCGGGCAACAAACACCGCTGGTAGCGGTGGTTTTT 2460
Db |||||
Db 1262 AAAAGAGTTGGTACTCTTGATTCGGGCAACAAACACCGCTGGTAGCGGTGGTTTTT 1203
Qy 2461 TGTTCGAGCAGCAGATTTACGCGCAGAAAAAGATCTCAAGAGATCTCTTCATCTT 2520
Db |||||
Db 1202 TGTTCGAGCAGCAGATTTACGCGCAGAAAAAGATCTCAAGAGATCTCTTCATCTT 1143
Qy 2521 TTCTACGGGGTCTGACGCTCAGTGGAAAGAAACTCAGGTTAAGGATTTTGGTCAATGAG 2580
Db |||||
Db 1142 TTCTACGGGGTCTGACGCTCAGTGGAAAGAAACTCAGGTTAAGGATTTTGGTCAATGAG 1083
Qy 2581 ATTATCAAAAAGGATCTTCACTAGATCTTTAAATTAATAATGAAGTTTAAATCAAT 2640
Db |||||
Db 1082 ATTATCAAAAAGGATCTTCACTAGATCTTTAAATTAATAATGAAGTTTAAATCAAT 1023
Qy 2641 CTAAAGTATATATGATGATAAATCTGTCTGACAGTTACCAATGCTTAATCAGTGGGCACC 2700
Db |||||
Db 1022 CTAAAGTATATATGATGATAAATCTGTCTGACAGTTACCAATGCTTAATCAGTGGGCACC 963
Qy 2701 TATCTCAGCGATCTCTATTTCTGTTTCCATCCATAGTTGCTGACTCCCGCTCGTGTAGAT 2760
Db |||||
Db 962 TATCTCAGCGATCTCTATTTCTGTTTCCATCCATAGTTGCTGACTCCCGCTCGTGTAGAT 903
Qy 2761 AACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGTGCAATGATACCGGAGACCC 2820
Db |||||
Db 902 AACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGTGCAATGATACCGGAGACCC 843
Qy 2821 ACGCTCACCGGCTCAGATTTTATCAGCAATAAACAGCAGCCGCGAAGCGCGAGCGAG 2880
Db |||||
Db 842 ACGCTCACCGGCTCAGATTTTATCAGCAATAAACAGCAGCCGCGAAGCGCGAGCGAG 783
Qy 2881 AAGTGGTCTCGCACTTTATCCGCTCCATCCAGTCTATTAATTTGTCGCGGAAGCTAG 2940
Db |||||
Db 782 AAGTGGTCTCGCACTTTATCCGCTCCATCCAGTCTATTAATTTGTCGCGGAAGCTAG 723
Qy 2941 AGTAAGTAGTTCCGCACTTAATAGTTTGGCGAA CGTTGTTGCCATTTGCTACAGGCATCGT 3000
Db |||||
Db 722 AGTAAGTAGTTCCGCACTTAATAGTTTGGCGAACTGTTGCCATTTGCTACAGGCATCGT 663
Qy 3001 GGTGTCAGCTCGTCTGTTGGTATGGCTTCAATTCAGCTCCGTTCCCAACGATCAAGGCG 3060
Db |||||
Db 662 GGTGTCAGCTCGTCTGTTGGTATGGCTTCAATTCAGCTCCGTTCCCAACGATCAAGGCG 603
Qy 3061 AGTTACATGATCCCCCATGTTGTGCAAAAAGCGGTTAGCTCTCTTCGCTCCCGATCGT 3120
Db |||||
Db 602 AGTTACATGATCCCCCATGTTGTGCAAAAAGCGGTTAGCTCTCTTCGCTCCCGATCGT 543
Qy 3121 TGTGGAAGTAGTTGGCGGAGGTATATCACTCATGTTTATGGCAGCAGCTGCAATAATTC 3180
Db |||||
Db 542 TGTGGAAGTAGTTGGCGGAGGTATATCACTCATGTTTATGGCAGCAGCTGCAATAATTC 483
Qy 3181 TCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGTTAGTACTCAACCAAGTC 3240
Db |||||
Db 482 TCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGTTAGTACTCAACCAAGTC 423
Qy 3241 ATTCTGAGAATAGTGTATCGGCGCAGCCGAGTTGCTCTTCGCGGCGGTCAATACCGGATAA 3300
Db |||||
Db 422 ATTCTGAGAATAGTGTATCGGCGCAGCCGAGTTGCTCTTCGCGGCGGTCAATACCGGATAA 363
Qy 3301 TACCGCGGCACATAGCAGAACTTTAAAGTGCTCATCATTTGGAAAAACGTTCTTCGGGGCG 3360
Db |||||
Db 362 TACCGCGGCACATAGCAGAACTTTAAAGTGCTCATCATTTGGAAAAACGTTCTTCGGGGCG 303

Qy 3361 AAACTCTCAAGGATCTTTACCGCTGTTGAGATCCAGTTTCAAGTGTAAACCCACTCGTGACC 3420
Db |||||
Db 302 AAACTCTCAAGGATCTTTACCGCTGTTGAGATCCAGTTTCAAGTGTAAACCCACTCGTGACC 243
Qy 3421 CAACTGATCTTCAGCATCTTTTACCTTTTACCAGCGTTTCCAGCGGTTTCTGGGTGAGCAAAACAGGAAG 3480
Db |||||
Db 242 CAACTGATCTTCAGCATCTTTTACCTTTTACCAGCGTTTCCAGCGGTTTCTGGGTGAGCAAAACAGGAAG 183
Qy 3481 GCAAAATCCGCAAAAAGGGAATTAAGGCGGACACGGAATGTTGAAATCTCATACTCTTT 3540
Db |||||
Db 182 GCAAAATCCGCAAAAAGGGAATTAAGGCGGACACGGAATGTTGAAATCTCATACTCTTT 123
Qy 3541 CCTTTTCAATATATTATGAAGCATTTATCAGGTTTATTGTCTCATGAGCGGATACATATT 3600
Db |||||
Db 122 CCTTTTCAATATATTATGAAGCATTTATCAGGTTTATTGTCTCATGAGCGGATACATATT 63
Qy 3601 TGAATGTTATTTAGAAAAATAAAACAAATAGGGGTTCCGCGCACATTTCCCGAAAAAGTGCC 3660
Db |||||
Qy 62 TGAATGTTATTTAGAAAAATAAAACAAATAGGGGTTCCGCGCACATTTCCCGAAAAAGTGCC 3
Qy 3661 AC 3662
Db ||
2 AC 1
RESULT 14
US-10-637-758-17/c
; Sequence 17, Application US/10637758
; Publication No. US20040180352A1
; GENERAL INFORMATION:
; APPLICANT: Large Scale Biology Corporation
; TITLE OF INVENTION: POPULATION OF POLYNUCLEOTIDE SEQUENCE VARIANTS
; FILE REFERENCE: LSBC-PLG5392-CIP
; CURRENT APPLICATION NUMBER: US/10/637,758
; PRIOR FILING DATE: 2003-08-08
; PRIOR APPLICATION NUMBER: US/10/280,913
; PRIOR FILING DATE: 2002-10-25
; PRIOR APPLICATION NUMBER: 60/402,342
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 10/066,390
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 60/268,785
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/266,386
; PRIOR FILING DATE: 2002-02-02
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 3637
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Encodes cycle 3 GFP
US-10-637-758-17
Query Match 72.8%; Score 2664.8; DB 18; Length 3637;
Best Local Similarity 83.8%; Pred. No. 0;
Matches 3070; Conservative 0; Mismatches 567; Indels 25; Gaps 4;
Qy 1 CTAAATGTGAAGCGTTAATAATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTC 60
Db 3637 CTAAATGTGAAGCGTTAATAATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTC 3578
Qy 61 ATTTTAAACCAATAGCGGAAATCGGCAAAATCCCTTATAAATCAAAAGATAGACCGA 120
Db 3577 ATTTTAAACCAATAGCGGAAATCGGCAAAATCCCTTATAAATCAAAAGATAGACCGA 3518
Qy 121 GATAGGGTTGAGTGTGTTGTCAGTTCAGTTCGAAACAGAGTCCACTATTAAAGACGTCGACTC 180
Db 3517 GATAGGGTTGAGTGTGTTGTCAGTTCGAAACAGAGTCCACTATTAAAGACGTCGACTC 3458
Qy 181 CAACGTCAAGCGGCAAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 240

|||||
3457 CAACGTCAAGGCGGAAACCGCTCTATCAGGGCGATGGCCCACTACGTGAACCATCAC 3398
QY
241 CTAATCAAGTTTTTGGGGT CGAGGTGCGGTAAAGCACTAAATCGGAACCCCTAAAGGGAG 300
|||||
3397 CTAATCAAGTTTTTGGGGT CGAGGTGCGGTAAAGCACTAAATCGGAACCCCTAAAGGGAG 3338
Db
301 CCGCCGATTTAGAGCTTGACGGGGAAAGCGGGGAAAGCTGGCGAGAAAGAGGGAAGAA 360
QY
3337 CCGCCGATTTAGAGCTTGACGGGGAAAGCGGGGAAAGCTGGCGAGAAAGAGGGAAGAA 3278
Db
361 AGCGAAGGAGCGGGCGCTGAGGCGCTGCAAGGTGAGCGGTCAACCTCGCGGTAAACAC 420
QY
3277 AGCGAAGGAGCGGGCGCTGAGGCGCTGCAAGGTGAGCGGTCAACCTCGCGGTAAACAC 3218
Db
421 CACACCGCGCGCTTAATTCGCGCGCTACAGGGCGGTCCCAATTCGCCATTCAGGCTGGG 480
QY
3217 CACACCGCGCGCTTAATTCGCGCGCTACAGGGCGGTCCCAATTCGCCATTCAGGCTGGG 3158
Db
481 CAACTGTTGGGAAGGGCGATCGGTGCGGCGCTTTCGCTATTACGCCAGCTGGCGGAAGG 540
QY
3157 CAACTGTTGGGAAGGGCGATCGGTGCGGCGCTTTCGCTATTACGCCAGCTGGCGGAAGG 3098
Db
541 GGGATGCTGCAAGGCGATTAAAGTTGGGTAAAGCGGAGGTTTCCCACTCAGGAGTTG 600
QY
3097 GGGATGCTGCAAGGCGATTAAAGTTGGGTAAAGCGGAGGTTTCCCACTCAGGAGTTG 3038
Db
601 TAAACGACGGCGAGTGGCGGCGCTTCATTCACGTTTTTGAACCGGTGGAGGAGCG 660
QY
3037 TAAACGACGGCGAGTGGCGGCGTAAAGTACGACTCACTATAGGGCGAATTTGAGCTCCA 2978
Db
661 GCAGACTCGCGGTGCAAAATGTTTTACAGCGTGAATGGAGCAGATGAAGATGCTCGACAC 720
QY
2977 CCGCGTGGCGG-----CGCTCTAGAACTAGTGGATCCCC--CGG 2939
Db
721 GCTGCAGACACGCGACTAGATTAAACCTAGAAAGATATCATATTGTGACGTACGTTAA 780
QY
2938 GCTGCAGGAATTTCTTATTTGTATAGTTTCATTCATGCCATGTGTAAATCCGACGACAGTTA 2879
Db
781 AGATAATCATGCGTAAATTTGACGATGGATCTGTAATACGACTCACTATAGGGCGAAT 840
QY
2878 CAAACTCAGAAGGACCATGTGGTCAGCTTTCGTTGGGATCTTTCGAAAGGGCAGATT 2819
Db
841 TGGGTACCGGGCCCCCTCGAGGTGCGAGGTATGATATGATATGATATCGAATTCCTGC 900
QY
2818 GTGTCGACAGGTAAATGTTGTCTGGTAAAGGACAGGGCCATCGCCAAATGGAGTATTTT 2759
Db
901 AGCCGGGGGATCCACTAGTTCTAGAGCGGCGCCACCGGGTGGAGCTCCAGCTTTTGT 960
QY
2758 GTTGATAATGGTCTGTAGTTGAACGGATCCATCTTCAATGTTGTGGCGAAATTTTGAAGT 2699
Db
961 TCCCTTTAGTGAGGGTTAATTAGATCCCATGGTCAATTTTACGCGAGACTATCTTTCTAG 1020
QY
2698 TAGCTTTGATCCATCTTTTGTGTTGCTGCGGTGATGATATCATTTGTGTGAG-TTATAG 2640
Db
1021 GGTAACTAGCTGATCAGGATCATATCGTGGGTCTTTTTTCGGGCTCAGTCACTCGCC 1080
QY
2639 TTGTACTCGAGTTTGTGTCGAGAAATGTTTCCATCTTCTTAAATCAATACCTTTTAAC 2580
Db
1081 CAACTGGCGCTATCTGGGCGATCGGGGAGGAGAACCGCGTCTTTTCCGGGAGGTG 1140
QY
2579 TCGATACGATTAAACAGGGTATCACCTTCAAACTTGACTTCAGCAAGCGCTTGTAGTTT 2520
Db
1141 AAGCGCATGGAAGAGTTTGGCGAGATGACTGCTGCTGATGATGAGTTGAGCGGAAC 1200
QY
2519 CCGTCATCTTTGAAGATATAGTGCCTCTCTGTACATAAACCTTCGGGCGATGGCACTCTTG 2460
Db
1201 GCACGTTTACCATGATGATTCGGGAAGGTGTGGCCATGCAAGCGCTTTAAACGGTGAAC 1260
QY
2459 AAAAGTCATCGCTTTTCATATGATCCGGATAACGGGAAAGCAATTTGAACACCATTAAG 2400
Db
1261 TCGTTACGCCACCTGGGATACAGTTTCGTGGCGCTTTTTCGGGACACAGTTTCCCGATGG 1320
QY

Db
2399 AAAGTAGTGACAAGTGTGTGGCCATGGAAACAGGTAGTATTTTCCAGTAGTGCAAAATAATTTA 2340
QY
1321 TCAGCCCGAAGCGCATCAGCAACCCGAAACAAATACCGCGACACCCGGAATCGCGTGCCG 1380
Db
2339 AGGGTAAGCTTTCCGTTATGTAGCATCACTTACCCCTCTCCACTGACAGAAAATTTGTGC 2280
QY
1381 GTGTGCAGATTATGACAGCGGTGCGGCGTGGGATATTACGTGAGGAGGAGGGGTATC 1440
Db
2279 CCATTAAACATCCATCTTAATTCAACAGAAATTTGGGACAACTCCAGTGAAGATTTCTTCT 2220
QY
1441 CTGGCTGGATGCGCAGAAATGGACATGATACCCCGTGAGTTACCCGGCGGCGCGCTT 1500
Db
2219 C---CTTTACTCATCGGTACCCAGCTTTTGTTCCTTTAGTGAGGGTTAATTGCGGCGCTT 2163
QY
1501 GCGTAATCATGCTCATAGCTGTTTCTGTGTAATTTGTTATCCGCTCAAAATTCACA 1560
Db
2162 GCGTAATCATGCTCATAGCTGTTTCTGTGTAATTTGTTATCCGCTCAAAATTCACA 2103
QY
1561 CAACTACGAGCGGAGCATAAAGTGTAAAGCTGGGTGGCTTAATGAGTGAGCTAACT 1620
Db
2102 CAACTACGAGCGGAGCATAAAGTGTAAAGCTGGGTGGCTTAATGAGTGAGCTAACT 2043
QY
1621 CACTTTAATTTGCTGCGCTCACTGCGCTTTCAGTCCGGAAAACCTGTGTGCCAGCT 1680
Db
2042 CACTTTAATTTGCTGCGCTCACTGCGCTTTCAGTCCGGAAAACCTGTGTGCCAGCT 1983
QY
1681 GCATTAATGAATCGGCAACCGCGGAGAGCGGTTTGGTATTTGGGCGCTTTCGCG 1740
Db
1982 GCATTAATGAATCGGCAACCGCGGAGAGCGGTTTGGTATTTGGGCGCTTTCGCG 1923
QY
1741 TTTCTGCTCATGCTGCTGCGCTCGGTTCGCTGCGGTGCGGCGAGCGGTATCAGTCA 1800
Db
1922 TTTCTGCTCATGCTGCTGCGCTCGGTTCGCTGCGGTGCGGCGAGCGGTATCAGTCA 1863
QY
1801 CTCAAAGCGGTAAATACGGTTATCCACAATCAGGGATAACGCGAGGAAGAACATGTG 1860
Db
1862 CTCAAAGCGGTAAATACGGTTATCCACAATCAGGGATAACGCGAGGAAGAACATGTG 1803
QY
1861 AGCAAAAGCCAGCAAAAGCCAGGAACCGTAAAGAGCGCGCTTCTGCGCTTTTTC 1920
Db
1802 AGCAAAAGCCAGCAAAAGCCAGGAACCGTAAAGAGCGCGCTTCTGCGCTTTTTC 1743
QY
1921 TAGGCTCCGCCCTCTGACGAGCATCAAAAATTCAGCGCTCAAGTCAAGGTGGGGA 1980
Db
1742 TAGGCTCCGCCCTCTGACGAGCATCAAAAATTCAGCGCTCAAGTCAAGGTGGGGA 1683
QY
1981 CCGCAGAGACTATAAAGATACAGCGCTTTCCTGGAAGCTCCCTGCTGCGCTCTCC 2040
Db
1682 CCGCAGAGACTATAAAGATACAGCGCTTTCCTGGAAGCTCCCTGCTGCGCTCTCC 1623
QY
2041 TGTTCGAGCCCTGCGCTTACCGGATACCTGTCGCGCTTTTCTCCCTTCGGGAAGCGTGGC 2100
Db
1622 TGTTCGAGCCCTGCGCTTACCGGATACCTGTCGCGCTTTTCTCCCTTCGGGAAGCGTGGC 1563
QY
2101 GCTTCTCATAGCTCACTGCTGATAGTATCTCAGTTCCGTTGAGGTGCTTCCGCTCAAGCT 2160
Db
1562 GCTTCTCATAGCTCACTGCTGATAGTATCTCAGTTCCGTTGAGGTGCTTCCGCTCAAGCT 1503
QY
2161 GGGCTGTGTGACGAAACCCCGCTTACGCGCTTTCAGCGCTTATCCGTTAACTATCG 2220
Db
1502 GGGCTGTGTGACGAAACCCCGCTTACGCGCTTTCAGCGCTTATCCGTTAACTATCG 1443
QY
2221 TCTGAGTCCAAACCGGTAAAGACACGACTTATCGCACCTGGCAGCAGCACTGGTAACAG 2280
Db
1442 TCTTGTGTTCAACCGGTAAAGACACGACTTATCGCACCTGGCAGCAGCACTGGTAACAG 1383
QY
2281 GATTAGCAGCGAGGTATGTAGGGGTGCTACAGATTTCTTGAAGTGTGGCTAACTA 2340
Db
1382 GATTAGCAGCGAGGTATGTAGGGGTGCTACAGATTTCTTGAAGTGTGGCTAACTA 1323
QY
2341 CCGCTACACTAGAGGAGAGTATTTGGTATCTGCGCTCTGCTGAGCCAGTTTACCTCGG 2400
Db
1322 CCGCTACACTAGAGGAGAGTATTTGGTATCTGCGCTCTGCTGAGCCAGTTTACCTCGG 1263

```
QY 2401 AAAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACACCCGCTGCTAGCGGTGGTTTTTT 2460
Db 1262 AAAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACACCCGCTGCTAGCGGTGGTTTTTT 1203
QY 2461 TGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTT 2520
Db 1202 TGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTT 1143
QY 2521 TTCTACGGGCTCAGCGCTCAGTGGAACGAAACCTCAGTTTAAGGATTTTGGTCATCAG 2580
Db 1142 TTCTACGGGCTCAGCGCTCAGTGGAACGAAACCTCAGTTTAAGGATTTTGGTCATCAG 1083
QY 2581 ATTATCAAAAAGGATCTTCACTAGATCCTTTTAAATTAATAATGAAGTTTAAATCAAT 2640
Db 1082 ATTATCAAAAAGGATCTTCACTAGATCCTTTTAAATTAATAATGAAGTTTAAATCAAT 1023
QY 2641 CTAAGATATATAGTAAACCTTGTGTCGACAGTTTACCAATGCTTAAATCAGTGAGGCACC 2700
Db 1022 CTAAGATATATAGTAAACCTTGTGTCGACAGTTTACCAATGCTTAAATCAGTGAGGCACC 963
QY 2701 TATCTCAGGGATCTGCTATTTGTTTCATCCATAGTTGCTGACTCCCGCTCGTGTAGT 2760
Db 962 TATCTCAGGGATCTGCTATTTGTTTCATCCATAGTTGCTGACTCCCGCTCGTGTAGT 903
QY 2761 AACTACGATACGGAGGGCTTACCATCTGGCCCCAGTGTGCAATGATACCGCAGACCC 2820
Db 902 AACTACGATACGGAGGGCTTACCATCTGGCCCCAGTGTGCAATGATACCGCAGACCC 843
QY 2821 ACGTCAACGGCTCAGATTATACAGCAATAAACACAGCAGCGGAAGGGCGAGCGAG 2880
Db 842 ACGTCAACGGCTCAGATTATACAGCAATAAACACAGCAGCGGAAGGGCGAGCGAG 783
QY 2881 AAGTGGTCTCGCACTTTATCGGCTCCATCCAGTCTATTAATTTGTCGGGAGCTAG 2940
Db 782 AAGTGGTCTCGCACTTTATCGGCTCCATCCAGTCTATTAATTTGTCGGGAGCTAG 723
QY 2941 AGTAAGTAGTTGCGCAGTTAATAGTTTGGCAACGTTGTGCAATTTGCTACAGGCATCGT 3000
Db 722 AGTAAGTAGTTGCGCAGTTAATAGTTTGGCAACGTTGTGCAATTTGCTACAGGCATCGT 663
QY 3001 GGTGTACGCTCGTGGTTGGTATGCTTCAATTCAGTCCGGTTCCTCAACGATCAAGGCG 3060
Db 662 GGTGTACGCTCGTGGTTGGTATGCTTCAATTCAGTCCGGTTCCTCAACGATCAAGGCG 603
QY 3061 AGTTACATGATCCCGATGTTGTGCABAAAGCGGTAGCTCCTTCGGTCCCTCCGATCGT 3120
Db 602 AGTTACATGATCCCGATGTTGTGCABAAAGCGGTAGCTCCTTCGGTCCCTCCGATCGT 543
QY 3121 TGTTCAGAAAGTAAAGTTGGCGCAGTGTATCACTCATGTTATGGCAGCACTGCATAATTC 3180
Db 542 TGTTCAGAAAGTAAAGTTGGCGCAGTGTATCACTCATGTTATGGCAGCACTGCATAATTC 483
QY 3181 TCTTACTGTATGCCATCCGTAAGATGCTTTTCTGTGACTGCTGAGTACTCAACCAAGTC 3240
Db 482 TCTTACTGTATGCCATCCGTAAGATGCTTTTCTGTGACTGCTGAGTACTCAACCAAGTC 423
QY 3241 ATTCTGAGAAATAGTATGCGCGCAGCCGATGCTCTTGGCCGGCGTCAATACGGGATAA 3300
Db 422 ATTCTGAGAAATAGTATGCGCGCAGCCGATGCTCTTGGCCGGCGTCAATACGGGATAA 363
QY 3301 TACCGCGCCACATACGAGAACTTTTAAAGTGCTCATCATTTGGAACCGTTCTTCGGGGCG 3360
Db 362 TACCGCGCCACATACGAGAACTTTTAAAGTGCTCATCATTTGGAACCGTTCTTCGGGGCG 303
QY 3361 AAAACTCTCAAGGATCTTTACCGCTGTTGAGATCCAGTTTCAATGTAACCACTCGTGCACC 3420
Db 302 AAAACTCTCAAGGATCTTTACCGCTGTTGAGATCCAGTTTCAATGTAACCACTCGTGCACC 243
QY 3421 CAACCTGATCTTCAGCATCTTTTACTTTTCCAGCGTCTTCTGGGTGAGCAAAACAGGAAG 3480
Db 242 CAACCTGATCTTCAGCATCTTTTACTTTTCCAGCGCTTCTTGGGTGAGCAAAACAGGAAG 183
```

```
QY 3481 GCAAAATCCCAAAAAAGGAAATAAGGCGCAGACGGAATGTTGAATACTCATCTCTT 3540
Db 182 GCAAAATCCCAAAAAAGGAAATAAGGCGCAGACGGAATGTTGAATACTCATCTCTT 123
QY 3541 CCTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGGGATACATATT 3600
Db 122 CCTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGGGATACATATT 63
QY 3601 TGAATGTATTTAGAAAAATAAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAAGTGCC 3660
Db 62 TGAATGTATTTAGAAAAATAAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAAGTGCC 3
QY 3661 AC 3662
Db 2 AC 1

RESULT 15
US-10-066-390-3/c
; Sequence 3, Application US/10066390
; Publication No. US20020146732A1
; GENERAL INFORMATION:
; APPLICANT: Padgett, Hal S.
; APPLICANT: Lindbo, John A.
; APPLICANT: Fitzmaurice, Wayne P.
; TITLE OF INVENTION: A Method of Increasing Complementarity
; FILE REFERENCE: P-LG 4878
; CURRENT APPLICATION NUMBER: US/10/066,390
; CURRENT FILING DATE: 2002-02-01
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 3637
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-10-066-390-3

Query Match 72.5%; Score 2656.4; DB 13; Length 3637;
Best Local Similarity 84.0%; Pred. No. 0;
Matches 3079; Conservative 0; Mismatches 556; Indels 29; Gaps 6;

QY 1 CTAATTCGTAAGCGTTAATATTTCGTTAAAAATTCGCGTTAAATTTTGTAAATCAGCTC 60
Db 3637 CTAATTCGTAAGCGTTAATATTTCGTTAAAAATTCGCGTTAAATTTTGTAAATCAGCTC 3578
QY 61 ATTTTAAACCAATAGGCGCAAAATCGGCAAAATCCCTTATATAATCAAAGAATAGACCGA 120
Db 3577 ATTTTAAACCAATAGGCGCAAAATCGGCAAAATCCCTTATATAATCAAAGAATAGACCGA 3518
QY 121 GATAGGTTGAGTGTGTTCCAGTTTGCAACAAGATGCCACTATTAAAGACGTGACCTC 180
Db 3517 GATAGGTTGAGTGTGTTCCAGTTTGCAACAAGATGCCACTATTAAAGACGTGACCTC 3458
QY 181 CAACGTCAAAGCGGCAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 240
Db 3457 CAACGTCAAAGCGGCAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 3398
QY 241 CTAATCAAGTTTTTTGGGGTCCAGGTGCGTAAAGCACTAAATCGGAACCCCTAAAGGGAG 300
Db 3397 CTAATCAAGTTTTTTGGGGTCCAGGTGCGTAAAGCACTAAATCGGAACCCCTAAAGGGAG 3338
QY 301 CCCCAGATTAGAGTTGACCGGGNAAGCCGCAACGTCGGGAGAAAGGAGGAGGAGAA 360
Db 3337 CCCCAGATTAGAGTTGACCGGGNAAGCCGCGAAGCGTGGCGAAGGAGGAGGAGAA 3278
QY 361 AGCGAAAGGAGCGGCGCTAGGCGCTGCGCAAGTGTAGCGGTACGCTGCGGTAAACCAC 420
Db 3277 AGCGAAAGGAGCGGCGCTAGGCGCTGCGCAAGTGTAGCGGTACGCTGCGGTAAACCAC 3218
QY 421 CACACCCCGCGCTTAAATAGCGCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCG 480
```

Db 3217 CACACCGCGCGCTTAATGCGCGCTACAGGGCGGTCCCATTCGCCAATTCAGGCTGCG 3158
Qy 481 CAACTGTTGGGAAGGCGATCGGTGCGGCTCTTCGCTATTACGCCAGCTGCGCGAAGG 540
Db 3157 CAACTGTTGGGAAGGCGATCGGTGCGGCTCTTCGCTATTACGCCAGCTGCGCGAAGG 3098
Qy 541 GGGATGTGTCGCAAGCGGATTAAGTTGGGTAAAGCGCAGGCTTTCCAGTCAGGCTTG 600
Db 3097 GGGATGTGTCGCAAGCGGATTAAGTTGGGTAAAGCGCAGGCTTTCCAGTCAGGCTTG 3038
Qy 601 TAAACGACGCGCAGTCAGCGGCTCTGCTTCAATTCAGCTTTTGAACCGCTGAGGACGG 660
Db 3037 TAAACGACGCGCAGTCAGCGGCTCTGCTTCAATTCAGCTTTTGAACCGCTGAGGACGG 2978
Qy 661 GCAGACTCGCGGTGCAAAATGTGTTTACAGCGTGTGAGGAGATGAAGATGCTCGACAC 720
Db 2977 CCGCGTGGCGCGCTCTAGAACTAGTGGATCCCGCGGCTGCAGGAATTCCTATTGTTGT 2918
Qy 721 GCTGCAGAACGCGAGCTAGATTAACTCTAGAAAGATTAATCATATGTCAGCTACGTTAA 780
Db 2917 ATAGTTTCAATCCATGCGCATGTGAATCCAGCAGCTGTTCAGAACTCAAGAGGACCATGT 2858
Qy 781 AGATAATC--ATGCGTAAATTTGACGATGGGATCTGTAATACGACTCACTATAGGCGA 838
Db 2857 GGTCTCTCTTTTGGGATCTTTTCGAAAGGCGAGATTGTGTGACAGGTAATGTTGT 2798
Qy 839 ATTGGGTACCGGCGCCCCCTCGAGGTCCAGCGTATCGATAAGCTTTGATATCGAATTCCT 898
Db 2797 CTGGTAAAGGACAGGCGCATGCCAAATGGAGTATTTTGTGATAATCGTCTGCTAGTT 2738
Qy 899 GCAGCGCGGGGATCCACTAGTTCTTAGAGCGGCGCCACCGCGGTGGAGCTCCAGCTTTT 958
Db 2737 GAAACGCTTCAATCTTCAATGTTGTCTAAATTTTGAAGTTAACTTTTGAATCCATCTTTT 2678
Qy 959 GTTCCCTTTAGTGGGTTAATTTAGATCCATGCGTCAATTTTACGACACTATCTTCT 1018
Db 2677 GTTGTCTGCCAATGATATCAATTTGTGTGA--GTATAGTTGATTTCCAAATTTGTGTCCA 2619
Qy 1019 AGGGTTAATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGCGCTCAGTCATCG 1078
Db 2618 AGAATGTTTCCATCTTCTTAAATCAATACCTTTTAACTCGATTC-----TATTAAACAA 2564
Qy 1079 CCCAAGCTGGCGCTATCTGGGCATCGGGAGGAAAGCGCGGTCTTTTCGCGGAGGT 1138
Db 2563 GGGTATCACTTCAAACTTGACTTC-----AGCACGTGCTTGTAGTTCCCGTC 2515
Qy 1139 TGAAGCGGATGGAAAGAGTTTGGCGAGGATGACTGCTGCTGCATTTGAGCTTTGAGCGRAA 1198
Db 2514 ATCCTTGAAAAATATAGTTCTTTCTGTATACATAACCTTCGGGCATGGCACTCTTGAAAAA 2455
Qy 1199 AGCGACGTTTACCATGATGATTTCGGGAAGGTGTGGCCATGCGCCTTTTAAACGGTGAAC 1258
Db 2454 GTCATGCGGTTTCAATATGATCTGGGTATCTTGAAAGCATTCGAAACCATATAGAAAGT 2395
Qy 1259 GTTCGTTACGCCACTCGGATACCGATTCTGCGGCTTTTCGAGACAGTTCCGGAT 1318
Db 2394 AGTGCAAGTGTGGCGATGGAACAGGTAGTTTTCAGTAGTGCAAAATAAATTTAAGGTT 2335
Qy 1319 GGTGAGCCGAGCGCATCAGCAACCCGNAACATACCGCGGACAGCCGGAATCGCGTGC 1378
Db 2334 AAGTTTTC-----CGTATGTCATCACCTTCACCTCTCCACTGACAGAAAAATTTGTGC 2280
Qy 1379 CGGTGTGCAGATTAAATGACAGCGGTGCGGCTGGGATATTACGTACAGCGAGCGGTA 1438
Db 2279 CCAATTAACATCACCATCTAATTTCAACAGAAATTTGGACAACTCCAGTGAAGTTCTTCT 2220
Qy 1439 TCTCGCTGGATGCGCGAGAAATGGACATGGATACCCGTTGAGTTACCGCGGCGCGC 1498
Db 2219 CCTTTACTCATCGGTACCCAGCTTTTGTTCCTTTAGTGAGGTTA-----ATTGCGCGC 2165
Qy 1499 TTGGGCTAATCATGGTCATAGCTGTTTCTGTGTGAATTTGTTATCCGCTCAATTTCA 1558

Db 2164 TTGGCGTAATCATGGTCATAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCACAAATCCA 2105
Qy 1559 CACAAATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAAATGAGTGAGCTAA 1618
Db 2104 CACAAATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAAATGAGTGAGCTAA 2045
Qy 1619 CTCACATTAATTTGGCTTGGCTCACCTGCGGCTTTTCCAGTCGGGAACCTGTGTCGCCAG 1678
Db 2044 CTCACATTAATTTGGCTTGGCTCACCTGCGGCTTTTCCAGTCGGGAACCTGTGTCGCCAG 1985
Qy 1679 CTGCAATTAATTTGGCCCAACGCGCGGGAGAGCGGTTTGGTATTTGGGCGCTCTTCC 1738
Db 1984 CTGCAATTAATTTGGCCCAACGCGCGGGAGAGCGGTTTGGTATTTGGGCGCTCTTCC 1925
Qy 1739 GCTTCTCTGCTCATGACTCGCTGCGCTCGGCTGCTGCGGTCGGGAGCGGTATCAGCT 1798
Db 1924 GCTTCTCTGCTCATGACTCGCTGCGCTCGGCTCGGCTCGGTCGGGAGCGGTATCAGCT 1865
Qy 1799 CACTCAAAAGCGGTAATACGTTATCCAGAAATCAGGGGATAACCGAGGAAGAACATG 1858
Db 1864 CACTCAAAAGCGGTAATACGTTATCCAGAAATCAGGGGATAACCGAGGAAGAACATG 1805
Qy 1859 TGAGCAAAAGGCGCAGCAAAAGGCGCAGGAACCGTAAAAAGCGCGCTTGTGGGCTTTTC 1918
Db 1804 TGAGCAAAAGGCGCAGCAAAAGGCGCAGGAACCGTAAAAAGCGCGCTTGTGGGCTTTTC 1745
Qy 1919 CATAGCTCCGCCCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGA 1978
Db 1744 CATAGCTCCGCCCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGA 1685
Qy 1979 AACCGACAGGACTATAAGATACAGGCGTTTCCCGCTTGGAAAGCTCCCTCGTCCGCTCT 2038
Db 1684 AACCGACAGGACTATAAGATACAGGCGTTTCCCGCTTGGAAAGCTCCCTCGTCCGCTCT 1625
Qy 2039 CCGTTCGACCGCTGCGCTTACCGGATACCTGTGCGCTTCTCTCCTTCGGGAAGCGTG 2098
Db 1624 CCGTTCGACCGCTGCGCTTACCGGATACCTGTGCGCTTCTCTCCTTCGGGAAGCGTG 1565
Qy 2099 GCGCTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTCTGTTGCTCCAAG 2158
Db 1564 GCGCTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTCTGTTGCTCCAAG 1505
Qy 2159 CTGGGCTGTGTCAGAAACCCCGTTACGCGGACCGCTGCGCTTATCCGGTAACTAT 2218
Db 1504 CTGGGCTGTGTCAGAAACCCCGTTACGCGGACCGCTGCGCTTATCCGGTAACTAT 1445
Qy 2219 CCGTTCGAGTCCAAACCGGTAAAGACGACTTATCGCCACTGCGACGACCTGCTGTAAC 2278
Db 1444 CCGTTCGAGTCCAAACCGGTAAAGACGACTTATCGCCACTGCGACGACCTGCTGTAAC 1385
Qy 2279 AGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTAAC 2338
Db 1384 AGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTAAC 1325
Qy 2339 TAGGCTACACTAGAAAGGACAGTATTTGGTATCTGGGCTCTGCTGAGGACGAGTTACCTTC 2398
Db 1324 TAGGCTACACTAGAAAGGACAGTATTTGGTATCTGGGCTCTGCTGAGGACGAGTTACCTTC 1265
Qy 2399 GGAAGAGAGTTGGTGTAGCTCTTGATCCGCAACCAACCCAGCTGAGCGTGGTGGT 2458
Db 1264 GGAAGAGAGTTGGTGTAGCTCTTGATCCGCAACCAACCCAGCTGAGCGTGGTGGT 1205
Qy 2459 TTTGTTTTCAGACGACAGATTACGCGCAGAAAAAAGGATCTCAAGAGATCTTTTGATC 2518
Db 1204 TTTGTTTTCAGACGACAGATTACGCGCAGAAAAAAGGATCTCAAGAGATCTTTTGATC 1145
Qy 2519 TTTTCTA GGGGTCTGACGCTCAGTGGAAACGAAAACTCAGTTAAGGGAATTTTGTCTATG 2578
Db 1144 TTTTCTA GGGGTCTGACGCTCAGTGGAAACGAAAACTCAGTTAAGGGAATTTTGTCTATG 1085
Qy 2579 AGATTATCAAAAGGATCTTCACTAGATCTTTTAAATTAATAATGAAGTTTAAATCA 2638
Db 1084 AGATTATCAAAAGGATCTTCACTAGATCTTTTAAATTAATAATGAAGTTTAAATCA 1025

Search completed: June 6, 2005, 04:28:09
Job time : 2022.86 secs

Qy	2639	ATCTAAAGATATATATAGATAAACTTGGTCTGACAGTTTACCAATGCTTAATCATAGTAGGCA	2698
Db	1024	ATCTAAAGATATATATAGATAAACTTGGTCTGACAGTTTACCAATGCTTAATCATAGTAGGCA	965
Qy	2699	CCTATCTCAGCGATCTGCTCTATTTTCGTTTCATCCATAGTTGCTGACTCCCGTCGTGTAG	2758
Db	964	CCTATCTCAGCGATCTGCTCTATTTTCGTTTCATCCATAGTTGCCTGACTCCCGTCGTGTAG	905
Qy	2759	ATAACTACGATACCGGAGGGCTTACCATCTTGCGCCCAAGTGTGCAATGATACCGCGAGAC	2818
Db	904	ATAACTACGATACCGGAGGGCTTACCATCTTGCGCCCAAGTGTGCAATGATACCGCGAGAC	845
Qy	2819	CCAGCTCACCGGCTCCAGATTTTATCAGCAATAAACACGACCGCGAAGGGCCGAGCGC	2878
Db	844	CCAGCTCACCGGCTCCAGATTTTATCAGCAATAAACACGACCGCGAAGGGCCGAGCGC	785
Qy	2879	AGAAAGTGGTCTCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTTGTCGCCGGGAAGCT	2938
Db	784	AGAAAGTGGTCTCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTTGTCGCCGGGAAGCT	725
Qy	2939	AGAGTAAGTAGTTTCGCCAGTTTAATAGTTTGGCGAAACGTTGTTGCCATTTGCTACAGGCATC	2998
Db	724	AGAGTAAGTAGTTTCGCCAGTTTAATAGTTTGGCGAAACGTTGTTGCCATTTGCTACAGGCATC	665
Qy	2999	GTGGTGTCAAGCTCGTTCGTTTGGTATAGCTTTCATTCAGCTCCGTTTCCCAACCATCAAGG	3058
Db	664	GTGGTGTCAAGCTCGTTCGTTTGGTATAGCTTTCATTCAGCTCCGTTTCCCAACCATCAAGG	605
Qy	3059	CGAGTTACATGATCCCCCATGTTGTGTGCAAAAAGCGTTAGTCTCTTCGGTCTCCGATC	3118
Db	604	CGAGTTACATGATCCCCCATGTTGTGTGCAAAAAGCGTTAGTCTCTTCGGTCTCCGATC	545
Qy	3119	GTTCTCAGAGTAAGTTGGCGCAGTGTATCACTCATGTGTATGGCAGCACATGCATAAT	3178
Db	544	GTTCTCAGAGTAAGTTGGCGCAGTGTATCACTCATGTGTATGGCAGCACATGCATAAT	485
Qy	3179	TCTCTTACTGTTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTAGTACTCAACCAAG	3238
Db	484	TCTCTTACTGTTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTAGTACTCAACCAAG	425
Qy	3239	TCATTCTGAGNATAGTGTATGCGCGCAGCCAGTTGCTCTTGCCCGCGGTCAATACGGGAT	3298
Db	424	TCATTCTGAGNATAGTGTATGCGCGCAGCCAGTTGCTCTTGCCCGCGGTCAATACGGGAT	365
Qy	3299	AATACCGCGCCACATAGCAGAACTTTTAAAAGTGTCTCATCTGGAAACGTTCTTCGGGG	3358
Db	364	AATACCGCGCCACATAGCAGAACTTTTAAAAGTGTCTCATCTGGAAACGTTCTTCGGGG	305
Qy	3359	CGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTTCGATGTAAACCACTCGTGCA	3418
Db	304	CGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTTCGATGTAAACCACTCGTGCA	245
Qy	3419	CCCAACTGATCTTCAGCATCTTTTACCTTCCACAGCGTTTCTCGGTGAGCAAAAACAGGA	3478
Db	244	CCCAACTGATCTTCAGCATCTTTTACCTTCCACAGCGTTTCTCGGTGAGCAAAAACAGGA	185
Qy	3479	AGGCAAAATGCCAAAAAGGGAATAAGGGCGACACGGAATAATGTTGAATACTCATACTC	3538
Db	184	AGGCAAAATGCCAAAAAGGGAATAAGGGCGACACGGAATAATGTTGAATACTCATACTC	125
Qy	3539	TTCCTTTTCAATATTATGAGCAATTTATCAGGGTTATTGTCTCATGAGCGGATACATA	3598
Db	124	TTCCTTTTCAATATTATGAGCAATTTATCAGGGTTATTGTCTCATGAGCGGATACATA	65
Qy	3599	TTTGAATGTATTATGAAAAATTAACAATAAGGGTTCGCGCGCATTTTCCCGGAAAGTG	3658
Db	64	TTTGAATGTATTATGAAAAATTAACAATAAGGGTTCGCGCGCATTTTCCCGGAAAGTG	5
Qy	3659	CCAC 3662	
Db	4	CCAC 1	

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 5, 2005, 13:37:08 ; Search time 15351.2 Seconds
(without alignments)
11558.877 Million cell updates/sec

Title: US-10-826-523-41

Perfect score: 3662

Sequence: 1 ctaattgtaagcgtaata.....atttccccgaaaagtgcac 3662

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sta.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
C 1	2664.8	72.8	3637	6	AX592968
C 2	2656.4	72.5	3637	6	AX592967
C 3	2650.6	72.4	3626	12	ASPGREEN1
C 4	2644.8	72.2	3624	12	XXU43957
C 5	2639	72.1	3681	6	A20702
C 6	2639	72.1	3681	6	I15651
C 7	2623.6	71.6	5314	6	A20700
C 8	2623.6	71.6	5314	6	I15649
C 9	2610.8	71.3	3552	12	AF171227
C 10	2609.4	71.3	4001	6	AX306330
C 11	2558.8	69.9	5540	12	AF179627
C 12	2550.6	69.7	3633	12	ASPGREEN2
C 13	2517.8	68.8	3918	6	CQ768840
C 14	2517.8	68.8	3918	6	CQ802124
C 15	2517.8	68.8	3918	6	CQ873225
C 16	2517.8	68.8	3918	6	CQ876882
C 17	2514.6	68.7	3886	12	XXU35129
C 18	2485.8	67.9	4141	12	AY190524
C 19	2484.6	67.8	3774	6	AX3063326

20	2474.2	67.6	5277	6	BD244458	BD244458	Genome an
21	2469.4	67.4	5277	6	AR340565	AR340565	Sequence
22	2451.8	67.0	3928	6	AX591027	AX591027	Sequence
23	2451.8	67.0	3928	6	AX591180	AX591180	Sequence
24	2451.8	67.0	3928	6	AX717603	AX717603	Sequence
25	2449.8	66.9	7933	12	AF359244	AF359244	Integrati
26	2448.8	66.9	3956	6	A84776	A84776	Sequence 21
27	2448.8	66.9	3956	6	AR300179	AR300179	Sequence
28	2448.8	66.9	3956	6	BD058066	BD058066	Method fo
29	2446.8	66.8	3715	6	AX306328	AX306328	Sequence
30	2434	66.5	3970	12	CVE403983	AJ403983	Cloning v
31	2420.6	66.1	3274	12	XXU43956	U43956	Expression
32	2414.6	65.9	3978	12	CVU41513	U41513	3' terminal
C 33	2401	65.6	3733	12	AF146395	AF146395	Cloning v
C 34	2396	65.4	3307	12	SYNPFABCX	D50399	Unidentifie
C 35	2382.6	65.1	4591	12	SYNPEZ218A	M74186	Cloning vec
C 36	2380.4	65.0	6624	12	CVU37458	U37458	Yeast CUP1
C 37	2372.2	64.8	4239	12	AVE401048	AJ401048	Artificia
38	2359	64.4	5781	12	CVU37457	U37457	Yeast CUP1
39	2357.4	64.4	4051	12	XXU35126	U35126	Plasmid pBS
C 40	2351	64.2	5275	12	CVU37456	U37456	Yeast CUP1
41	2349.8	64.2	3230	12	AY663843	AY663843	Cloning v
42	2340.2	63.9	3230	12	AY663844	AY663844	Cloning v
C 43	2333.4	63.7	5436	12	AY303166	AY303166	CFP Integ
C 44	2332.8	63.7	5436	12	AY303167	AY303167	YFP Integ
45	2328.6	63.6	3204	12	AY663845	AY663845	Cloning v

ALIGNMENTS

RESULT 1	AX592968/c	AX592968	3637 bp	DNA	linear	PAT 27-JAN-2003
LOCUS	Sequence 4 from Patent WO02079468.					
DEFINITION	AX592968					
ACCESSION	AX592968.1	GI:27950812				
VERSION						
KEYWORDS	synthetic construct					
SOURCE	synthetic construct					
ORGANISM	other sequences; artificial sequences.					
REFERENCE	1					
AUTHORS	Padgett,H.S., Fitzmaurice,W.P. and Lindo,J.A.					
TITLE	A method of increasing complementarity in a heteroduplex					
JOURNAL	Patent: WO 02079468-A 4 10-OCT-2002;					
	Large Scale Biology Corporation (US)					
FEATURES	Location/Qualifiers					
source	1..3637					
	/organism="synthetic construct"					
	/mol_type="unassigned DNA"					
	/db_xref="taxon:32630"					

72.8%; Score 2664.8; DB 6; Length 3637;						
Best Local Similarity 83.8%; Pred. No. 0;						
Matches 3070; Conservative 0; Mismatches 567; Indels 25; Gaps 4;						
Qy	1	CTAAATTGTAAGCGTTAAATATTTTGTAAAAATTCGCGTTAAATTTTGTAAATCAGCTC	60			
Db	3637	CTAAATTGTAAGCGTTAAATATTTTGTAAAAATTCGCGTTAAATTTTGTAAATCAGCTC	3578			
Qy	61	ATTTTTTAAACCAATAGCGCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA	120			
Db	3577	ATTTTTTAAACCAATAGCGCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA	3518			
Qy	121	GATAGCGTTGAGTGTGTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGAACTGGACTC	180			
Db	3517	GATAGCGTTGAGTGTGTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGAACTGGACTC	3458			
Qy	181	CAACGTCAAAGGCGCAAAAACCGTCTATCAGGCGGATGCCCACTACGTGAACCATCACC	240			
Db	3457	CAACGTCAAAGGCGCAAAAACCGTCTATCAGGCGGATGCCCACTACGTGAACCATCACC	3398			

QY 241 CTAATCAAGTTTTTTGGGTCGAGTGCCTAAAGCACTAAATCGGAACCCCTAAAGGGAG 300
DB |||||
3397 CTAATCAAGTTTTTTGGGTCGAGTGCCTAAAGCACTAAATCGGAACCCCTAAAGGGAG 3338
QY 301 CCCCGAATTTAGAGCTTTGACGGGGAAAGCCCGGGAACGTTGGCGAGAAAGGAAGGAAGAA 360
DB |||||
3337 CCCCGAATTTAGAGCTTTGACGGGGAAAGCCCGGGAACGTTGGCGAGAAAGGAAGGAAGAA 3278
QY 361 AGCGAAAGAGCGGCGCTAGGCGCTGCAAGTGTAGCGGTACGCTGCGGCTAACAC 420
DB |||||
3277 AGCGAAAGAGCGGCGCTAGGCGCTGCAAGTGTAGCGGTACGCTGCGGCTAACAC 3218
QY 421 CACACCGCGCGCTTAATGCGCGCGCTACAGGGCGCTCCCATTCGCCATTACAGCTGCG 480
DB |||||
3217 CACACCGCGCGCTTAATGCGCGCGCTACAGGGCGCTCCCATTCGCCATTACAGCTGCG 3158
QY 481 CAACTGTTGGGAAGGGCGAGTCGGTGGGCGCTCTTCGCTATTACGCGAGCTGGCGAAAGG 540
DB |||||
3157 CAACTGTTGGGAAGGGCGAGTCGGTGGGCGCTCTTCGCTATTACGCGAGCTGGCGAAAGG 3098
QY 541 GGGATGTCGCAAGCGCAATTAAAGTTGGTAACGCCAGGTTTTCCCACTACAGCACTTG 600
DB |||||
3097 GGGATGTCGCAAGCGCAATTAAAGTTGGTAACGCCAGGTTTTCCCACTACAGCACTTG 3038
QY 601 TAAACGAGCGGCAAGTGAAGCGCGCTCGTTCAATTCACGTTTTTGAACCCGTCGAGGACGG 660
DB |||||
3037 TAAACGAGCGGCAAGTGAAGCGCGCTCAATACGACTCACTATAGGGCGAATTGGAGCTCA 2978
QY 661 GCAGACTCGCGTGCAATGTGTTTTACAGCGTGTATGGAGCAGATGAAGATGCTCGACAC 720
DB |||||
2977 CCGCGTGGCGG-----CCGCTCTAGAACTAGTGATCCCG--CGG 2939
QY 721 GCTGCAGAACCGCAGCTAGATTAAACCTTAGAAAGATATCATATTGTGACGTACGTTAA 780
DB |||||
2938 GCTGCAGGAATTCATTATTGTATGATTCATCATGCCATGTGTAATCCAGCAGCAGTTA 2879
QY 781 AGATAATCATGCGTAAAAATTGACGATGGGATCTGTAAATACGACTCACTATAGGGCGAAT 840
DB |||||
2878 CAAACTCAAGAGGACCATGTGGTCACGCTTTTGGTGGGATCTTTCGAAAGGGCAGATT 2819
QY 841 TGGGTACCGGGCCCCCTCGAGGTCGAGGTATCGAATAGCTTGAATCGAATTCCTGCG 900
DB |||||
2818 GTGTCGACAGGTAAATGTTGTCTGTGTAAAGGACAGGGCCATCGCCAAATTGGAGTATTTT 2759
QY 901 AGCCCCGGGGATCCACTAGTTCTAGAGCGCGCCACCGCGTGGAGCTCCAGCTTTTGT 960
DB |||||
2758 GTTGATATGCTCTGCTAGTTGAACGGATCCATCTTCAATGTTGTGGCAATTTTGAAGT 2699
QY 961 TCCCTTTAGTGAAGGTAAATTAGATCCCATGCGTCAATTTTACGAGACTATCTTTCTAG 1020
DB |||||
2698 TAGCTTTGATTCATTTTGTCTGCGGTGATGTATACATTTGTGTGAG-TTATAG 2640
QY 1021 GGTAACTAGCTGATCAGGATCATATCGTGGGTCTTTTCCGGCTCAGTCACTGCG 1080
DB |||||
2639 TTGTACTCGAGTTTGTCTCCGAGAATGTTTCCATCTCTTTAAAAATCAATACCTTTTAA 2580
QY 1081 CAACTGCGGCTATCTGGCATCGGGAGGAAGACCGCTGCTTTTCCGCGAGGTTG 1140
DB |||||
2579 TCGATACGATTAACAGGGTATACCTTCAAACTTGACCTTACGACGCGTCTTGTAGTTC 2520
QY 1141 AAGCGGCATGGAAGAGTTTTCGAGGATGACTGCTGTGATGACGTTGAGCGTGAAGAAAC 1200
DB |||||
2519 CCGTCATCTTTGAAAGATATAGTGGTCTCTGTACATAAACCTTCGGGCATGGCACTCTG 2460
QY 1201 GCACGTTTACCATGATGATTCGGGAAGGTGTGGCCATGACGCGCTTTAAACGCTGAATGT 1260
DB |||||
2459 AAAAAAGTCATGCGCTTTCATATGATCGGATTAACGGGAAAAGCATTTGAACACCAATAAG 2400
QY 1261 TCGTTACGCCACCTGGGATACAGTTGTCGCGCTTTTCCGGACACAGTTTCCGGATGG 1320
DB |||||
2399 AAAGTAGTGACAAAGTTTGGCCATCGAACAGGTAGTTTTCAGTAGTGCMAATAAATTA 2340
QY 1321 TAGCCCCGAAGCGCATACGCAACCCGAAACAATACCGGGCAGACCGGAACCTGCGTGGCG 1380

DB 2339 AGGGTAAGCTTTTCGCTATGTAGCATCACCTTCACTCTCCACTGACAGAAAATTTGTGC 2280
QY 1381 GTGTGCAGATTAATACAGCGGTGGCGCTGGGATATATCGTCAAGCAGGAGGGGTATC 1440
DB 2279 CCATTTAAATCATCACTACCATTAATTTCAAACAAGAAATGGGACAACTCCAGTGAAGATTTCTTCT 2220
QY 1441 CTGGCTGATGCCGAGAAATGGAATGGAATGATGATACCCCGTGAGTTACCCGCGCGGCGCTT 1500
DB 2219 C---CTTTACTCATCGGTACCCAGCTTTTGTTCCTTTAGTGAGGGTTAATTTGCGGCTT 2163
QY 1501 GCGCTAATCATGGTCATAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCACAAATTTCCACA 1560
DB 2162 GCGCTAATCATGGTCATAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCACAAATTTCCACA 2103
QY 1561 CAACTATGAGCGCGGAAGCATAAAGTGAAGCCTGGGCTGCTAATGATGAGCTAATCT 1620
DB 2102 CAACTATGAGCGCGGAAGCATAAAGTGAAGCCTGGGCTGCTAATGATGAGCTAATCT 2043
QY 1621 CAACTAATTTGGTTCGCTCACTCCGCTTCCAGTCCGGAACCTGTCTGTCGAGCT 1680
DB 2042 CAACTAATTTGGTTCGCTCACTCCGCTTCCAGTCCGGAACCTGTCTGTCGAGCT 1983
QY 1681 GAACTAATGAAATCGCCCAAACGCGCGGGAGAGCGGTTTGCCTATTTGGGCGCTCTTCCGC 1740
DB 1982 GAACTAATGAAATCGCCCAAACGCGCGGGAGAGCGGTTTGCCTATTTGGGCGCTCTTCCGC 1923
QY 1741 TTCCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCA 1800
DB 1922 TTCCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCA 1863
QY 1801 CTCGAGGCGGTAATACGCTTATCCAGAAATCAGGGGATTAACGAGGAGAAACAATGTG 1860
DB 1862 CTCGAGGCGGTAATACGCTTATCCAGAAATCAGGGGATTAACGAGGAGAAACAATGTG 1803
QY 1861 AGCAAAAGCGCAGCAAAAAGGCGGAGAACCGTAAAGGCGCGGTTGCTGGCGTTTTCCTCA 1920
DB 1802 AGCAAAAGCGCAGCAAAAAGGCGGAGAACCGTAAAGGCGCGGTTGCTGGCGTTTTCCTCA 1743
QY 1921 TAGGCTCGCGCCCTGAGAGCATCAAAAATCAGCGCTCAAGTCAGAGGTGGCGAA 1980
DB 1742 TAGGCTCGCGCCCTGAGAGCATCAAAAATCAGCGCTCAAGTCAGAGGTGGCGAA 1683
QY 1981 CCCGACAGGACTATAAAGATACAGGCGTTTCCCTCTGGAAGCTCCCTGTCGCTCTCC 2040
DB 1682 CCCGACAGGACTATAAAGATACAGGCGTTTCCCTCTGGAAGCTCCCTGTCGCTCTCC 1623
QY 2041 TGTTCGAGCCCTGCGCTTACCGGATACCTGTCCGCTTTTCCCTTCGGGAAGCGTGGC 2100
DB 1622 TGTTCGAGCCCTGCGCTTACCGGATACCTGTCCGCTTTTCTCCCTTCGGGAAGCGTGGC 1563
QY 2101 GCTTTCATAGCTCACGCTGATGATCTCAGTTTGGGTAGTGTCTGCTTCCCAAGCT 2160
DB 1562 GCTTTCATAGCTCACGCTGATGATCTCAGTTTGGGTAGTGTCTGCTTCCCAAGCT 1503
QY 2161 GGGCTGTGTGACGAAACCCCGTTTCAGCGCGAGCGCTGCGCTTTATCCGCTTAATCTCG 2220
DB 1502 GGGCTGTGTGACGAAACCCCGTTTCAGCGCGAGCGCTGCGCTTTATCCGCTTAATCTCG 1443
QY 2221 TCTTGAAGTCCAAACCGGTTAAGCAGCATTTACGCACTTGGCAGCAGCCTGTTAAACG 2280
DB 1442 TCTTGAAGTCCAAACCGGTTAAGCAGCATTTACGCACTTGGCAGCAGCCTGTTAAACG 1383
QY 2281 GATTAGCAGCGAGGTATGAGGGGTGCTACAGATTCTTGAAGTGGTGGCTTAACCTA 2340
DB 1382 GATTAGCAGCGAGGTATGAGGGGTGCTACAGATTCTTGAAGTGGTGGCTTAACCTA 1323
QY 2341 CCGCTTACACTAGAAGCAGATTTTGGTATCTGCGCTCTGCTGTAAGCAGCTTACCTCGG 2400
DB 1322 CCGCTTACACTAGAAGCAGATTTTGGTATCTGCGCTCTGCTGTAAGCAGCTTACCTCGG 1263
QY 2401 AAAAGAGTTGGTAGCTTGTGATCCGGCAACAAACCAACCGCTGCTGAGGGTGGTTTTT 2460

Db 1262 AAAAAAGTTGGTAGCTCTTGATCCGGCAAAACAAACACACCGCTGGTAGCGGTGTTTTT 1203
QY 2461 TGTTCAGACGACAGATTACGGCGAGAAAAAGGATCTCAAGAAAGATCCTTTTGATCTT 2520
Db 1202 TGTTTCAAGCAGCAGATTACGGCGAGAAAAAGGATCTCAAGAAAGATCCTTTTGATCTT 1143
QY 2521 TTCTACGGGTCTGACGCTCAGTGGAAACGAAACTCAGTTAAGGATTTTGTCATGAG 2580
Db 1142 TTCTACGGGTCTGACGCTCAGTGGAAACGAAACTCAGTTAAGGATTTTGTCATGAG 1083
QY 2581 ATTATCAAAAAGGATCTTCACTCAGATCTCTTTTAAATTAATAATGAAGTTTAAATCAAT 2640
Db 1082 ATTATCAAAAAGGATCTTCACTCAGATCTCTTTTAAATTAATAATGAAGTTTAAATCAAT 1023
QY 2641 CTAAAGTATATATGAGTAAACTTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACC 2700
Db 1022 CTAAAGTATATATGAGTAAACTTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACC 963
QY 2701 TATCTCAGGATCTGTCTATTTCGTTTCATTCATGATAGTTGCTGACTCCCGTCTGTAGAT 2760
Db 962 TATCTCAGGATCTGTCTATTTCGTTTCATTCATGATAGTTGCTGACTCCCGTCTGTAGAT 903
QY 2761 AACTCAGATACGGAGGGCTTACCATCTGSCCCAGTGTGCAATGATACCGGAGACCC 2820
Db 902 AACTCAGATACGGAGGGCTTACCATCTGSCCCAGTGTGCAATGATACCGGAGACCC 843
QY 2821 ACCTCACCAGGCTCCAGATTTATCAGCAATAAACACGACCGAGGCGAGGCGAG 2880
Db 842 ACCTCACCAGGCTCCAGATTTATCAGCAATAAACACGACCGAGGCGAGGCGAG 783
QY 2881 AGTGTCTCTGAACTTTATCCGCTCCATCCAGTCTATTAATGTTGCGGGAAGCTAG 2940
Db 782 AGTGTCTCTGAACTTTATCCGCTCCATCCAGTCTATTAATGTTGCGGGAAGCTAG 723
QY 2941 AGTAACTAGTTCGCCAGTTAATAGTTTGGCAAGTTTGTGCCATGCTACAGGCATCGT 3000
Db 722 AGTAACTAGTTCGCCAGTTAATAGTTTGGCAAGTTTGTGCCATGCTACAGGCATCGT 663
QY 3001 GGTGTCAAGCTCGTCTGTTGGTATGCTTCATTCAGCTCCGGTTCACCAAGCATCAAGGCG 3060
Db 662 GGTGTCAAGCTCGTCTGTTGGTATGCTTCATTCAGCTCCGGTTCACCAAGCATCAAGGCG 603
QY 3061 AGTTACATGATCCCAATGTTGGTGGAAAGGCGTTAGCTCTTCGCTCCCGATCGT 3120
Db 602 AGTTACATGATCCCAATGTTGGTGGAAAGGCGTTAGCTCTTCGCTCCCGATCGT 543
QY 3121 TGTCAAGTAAAGTTCGCCAGTGTATCAGTTCATGTTATGCGAGCACTGCATATTC 3180
Db 542 TGTCAAGTAAAGTTCGCCAGTGTATCAGTTCATGTTATGCGAGCACTGCATATTC 483
QY 3181 TCTTACTGTCAAGCTCCGTAAGATGCTTTTCTGTGACTGGTGGTGAAGTCAACCAAGTC 3240
Db 482 TCTTACTGTCAAGCTCCGTAAGATGCTTTTCTGTGACTGGTGGTGAAGTCAACCAAGTC 423
QY 3241 ATTCTGAGAAATAGTGTATGCGGACCGAGTTGCTCTTGCCCGGCTCAATACGGGATAA 3300
Db 422 ATTCTGAGAAATAGTGTATGCGGACCGAGTTGCTCTTGCCCGGCTCAATACGGGATAA 363
QY 3301 TACCGGCGCACATAGCAGAACTTTAAAGTCTCATCATTTGGAAGAAAGTCTTCGCGGCG 3360
Db 362 TACCGGCGCACATAGCAGAACTTTAAAGTCTCATCATTTGGAAGAAAGTCTTCGCGGCG 303
QY 3361 AAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCAGATGTAACCACTCGTGACC 3420
Db 302 AAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCAGTGAACCACTCGTGACC 243
QY 3421 CAACTGATCTTACGATCTTTTACTTTCACAGCGTTTTCAGCGGTTTTCGGGTGAGCAAAAACAGGAAG 3480
Db 242 CAACTGATCTTACGATCTTTTACTTTCACAGCGTTTTCGGGTGAGCAAAAACAGGAAG 183
QY 3481 GCAAAATGCGCAAAAAGGGAATAGGGGACACCGGAATGTTGAATCTCATCTCTT 3540
Db 182 GCAAAAATGCGCAAAAAGGGAATAGGGGACACCGGAATGTTGAATCTCATCTCTT 123

QY 3541 CCTTTTCAATATATTATGAAGCAATTTATCAGGGTTATTTGCTCATGAGCGGATACATATT 3600
Db 122 CCTTTTCAATATATTATGAAGCAATTTATCAGGGTTATTTGCTCATGAGCGGATACATATT 63
QY 3601 TGAATGTATTTAGAAAAATAACAAATAGGGTTCCGCGCACATTTCCCGAAAAAGTGCC 3660
Db 62 TGAATGTATTTAGAAAAATAACAAATAGGGTTCCGCGCACATTTCCCGAAAAAGTGCC 3
QY 3661 AC 3662
Db 2 AC 1
RESULT 2
LOCUS AX592967/c 3637 bp DNA linear PAT 27-JAN-2003
DEFINITION Sequence 3 from Patent WO02079468.
ACCESSION AX592967
VERSION AX592967.1 GI:27950811
KEYWORDS synthetic construct
SOURCE other sequences; artificial sequences.
ORGANISM Padgett,H.S., Fitzmaurice,W.P. and Lindo,J.A.
REFERENCE 1
AUTHORS A method of increasing complementarity in a heteroduplex
TITLE Patent: WO 02079468-A 3 10-OCT-2002;
JOURNAL Large Scale Biology Corporation (US)
FEATURES Location/Qualifiers
source 1. .3637
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
ORIGIN
Query Match 72.5%; Score 2656.4; DB 6; Length 3637;
Best Local Similarity 84.0%; Pred. No. 0;
Matches 3079; Conservative 0; Mismatches 556; Indels 29; Gaps 6;
QY 1 CTAAATGTGAAGCGTTAATAATTTTGTAAATTCGGGTTAAATTTTGTAAATCAGCTC 60
Db 3637 CTAAATGTGAAGCGTTAATAATTTTGTAAATTCGGGTTAAATTTTGTAAATCAGCTC 3578
QY 61 ATTTTAAACCAATAGCGCGAAATCGCAAAATCCCTTATAAATCAAAAGAAATAGACCGA 120
Db 3577 ATTTTAAACCAATAGCGCGAAATCGCAAAATCCCTTATAAATCAAAAGAAATAGACCGA 3518
QY 121 GATAGGTTGAGTGTGTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGAACGTGGACTC 180
Db 3517 GATAGGTTGAGTGTGTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGAACGTGGACTC 3458
QY 181 CAACTGCAAGGCGGCAAAACCGTCTATCAGGGCGATGGCCCACTAGTGAACCATCACC 240
Db 3457 CAACTGCAAGGCGGCAAAACCGTCTATCAGGGCGATGGCCCACTAGTGAACCATCACC 3398
QY 241 CTAATCAAGTGTTCGGGTCGAGGTCCGTAAGCACTAAATCGGAACCTTAAGGGAG 300
Db 3397 CTAATCAAGTGTTCGGGTCGAGGTCCGTAAGCACTAAATCGGAACCTTAAGGGAG 3338
QY 301 CCCCCGATTTAGAGCTTGACCGGGAAAGCCGGCGAAACGTGGCGAGAAAGGAAGAAAGAA 360
Db 3337 CCCCCGATTTAGAGCTTGACCGGGAAAGCCGGCGAAACGTGGCGAGAAAGGAAGAAAGAA 3278
QY 361 AGCGAAGGAGCGGGCGCTAGGGCGCTGACAGTGTAGCGGTCAAGTCCGCTGCGGTAAACAC 420
Db 3277 AGCGAAGGAGCGGGCGCTAGGGCGCTGACAGTGTAGCGGTCAAGTCCGCTGCGGTAAACAC 3218
QY 421 CACACCCGCGCGCTTAAATGCGCGCTACAGGGCGGCTCCCATTCGCAATTCAGGCTGCG 480
Db 3217 CACACCCGCGCGCTTAAATGCGCGCTACAGGGCGGCTCCCATTCGCAATTCAGGCTGCG 3158
QY 481 CAACTGTGGGAAGGGCGGATCGGTGCGGCGCTCTTCGCTATTACGCCAGCTGGCGAAAGG 540

Db 3157 CAACTGTTGGGAAGGCGCATCGGTGCGGCCCTCTTCGCTATTATACGCCAGCTGGCGAAAGG 3098
QY 541 GGGATGCTCTGCAAGGCGATTAAAGTTGGGTACGCCAGGGTTTTCCCACTCAGCAGCTTG 600
Db 3097 GGGATGCTGCTCAAGGCGCATTAAGTTGGGTAAACGCCAGGGTTTTCCCACTCAGCAGCTTG 3038
QY 601 TAAAAACGAGCGGCAGTGTAGCGCGCTCGTTCAATTCACGTTTTTGAAACCCGTGGAGGAGG 660
Db 3037 TAAAAACGAGCGGCAGTGTAGCGCGCTTAATACGACTCACTATAGGCGGAATTGGAGCTCCA 2978
QY 661 GCAGACTCGCGTGCAAATGTTTTACAGCTGTGATGGAGCAGATGAAGATGCTCGACAC 720
Db 2977 CCGCGTGGCGGCCGCTCTAGAACTAGTGTATCCCGCGCTGCAGGAATCTTATTGTT 2918
QY 721 GCTGCAGAACGCGAGCTAGATTAAACCTTAGAAGATTAATCATATATGTCACCTAGTTAA 780
Db 2917 ATAGTTCAATCCATGCCATGTGTAAATCCCGAGCGTGTACAAACTCAAGAAGGACATGT 2858
QY 781 AGATAATC--ATGCGTAAATTTGACGCATGGGATCTGTAATACGACTCACTATAGGGGA 838
Db 2857 GGTCTCTCTTTTCGTTGGGATCTTTGAAAGGCGAGATTGTGTGACAGGTAATGGTGT 2798
QY 839 ATTGGGTACCGGGCCCCCTCGAGGTGAGCGGTATCGATAAGCTTGATATCGAATTCCT 898
Db 2797 CTGGTAAAGGACAGGGCCATCGCCAAATTGGAGTATTTGTTGATAATGGTCTGCTAGTT 2738
QY 899 GCAGCCGGGGGATCCACTAGTTCTTAGAGCGGCCCCAGCGGGTGGAGCTCCAGCTTTT 958
Db 2737 GAACGCTTCCATCTTCAATGTTGTGCTAAATTTTGAAGTTAACTTTGATTCATCTTTT 2678
QY 959 GTTCCCTTTAGTGGGTAAATTTAGATCCATGCGTCAATTTTACGACACTATCTTTCT 1018
Db 2677 GTTGTCTGCCATGATATACATTTGTGTGA--GTTATAGTTGATTTCCAAATTGTGTCCA 2619
QY 1019 AGGGTAAATCTAGCTGCATCAGGATCATATGTCGGGTCTTTTTTCGGGCTCAGTCACTG 1078
Db 2618 AGAATGTTTCCATCTCTTTAAATCAATACCTTTTAACTCGATTC-----TATTAAACA 2564
QY 1079 CCCAAGCTGGGCTATCTGGGATCGGGGAGGAGAGCCGTCGCTTTCCGCGAGGT 1138
Db 2563 GGGTATACCTTCAAACTTGACTTC-----AGCACGTGCTCTTGATGTTCCCGTC 2515
QY 1139 TGAAGCGGCATGGAAGAGTTTGGCGAGGATGACTGCTGCTCAATTGACGTTGAGCGGAA 1198
Db 2514 ATCCTTGAATAATATAGTTCTTCTGTACATAAACCCTTCGGGCATGSCACTCTTGAANA 2455
QY 1199 AGCAGCTTACCATGATGATTCGGGAAGGTGTGCCATGCAAGCCTTTTAAAGGTGAAC 1258
Db 2454 GTCATGCGGTTTCATATGATCTGGGTATCTTGAAGAAGCAATTGAACACCATAGAGAAAGT 2395
QY 1259 GTTCGTTCAGGCCACTGGGATACAGTTGCTGCGGGCTTTTCCGGACACAGTTCGGAT 1318
Db 2394 AGTGACAAGTGTGTGCCATGGAAACAGGTAGTTTTCAGTAGTGGCAATAAATTTAAGGT 2335
QY 1319 GGTGAGCCGAGCGCATCAGCAACCCGAAATAATACCGCGCAGCAGCGGAACTGCGCTGC 1378
Db 2334 AAGTTTTC-----CGTATGTTGCATCACCTTCACTTCCCTCTCCACTGACAGANAATTGTGC 2280
QY 1379 CGGTGTGAGATTAATGACAGCGGTGCGCGCTGGGATATTAACGTGACGAGGACGGGTA 1438
Db 2279 CCATTAAACATCACCATCTAATTTCAACAAGAAATGGGACAACTCCAGTGAAGATTTCTCT 2220
QY 1439 TCCTGGCTGGATGCCGAGANAATGACATGGATACCCGTAGTTTACCGCGGGCGGC 1498
Db 2219 CCTTTACTCATCGGTACCGAGCTTTTGTTCCTTTTAGTGAGGGTTA-----ATTGCGGC 2165
QY 1499 TTGGGTAATCATGCTCATAGCTGTTTCTGTGTGAATTTGTTATCCGCTCACAAATTTCCA 1558
Db 2164 TTGGGTAATCATGCTCATAGCTGTTTCTGTGTGAATTTGTTATCCGCTCACAAATTTCCA 2105
QY 1559 CACAACTACGAGCGGAGGCATAAAGGTAAAGCTGGGGTGCCTTAATGATGAGCTAA 1618
|||||

Db 2104 CACAACTACGAGCGGAGGCATAAAGTGTAAAGCCTGGGGTGCCCTAATGATGAGCTAA 2045
QY 1619 CTCACATTAATTTGCTGCGCTCACTGCGCGCTTTCAGTTCGGGAAACCTGTCGTGCCAG 1678
Db 2044 CTCACATTAATTTGCTGCGCTCACTGCGCGCTTTCAGTTCGGGAAACCTGTCGTGCCAG 1985
QY 1679 CTGCATTAATGAATCGGCCAACCGCGGGGAGAGCGGTTGCGTATTGGGCGCTCTTCC 1738
Db 1984 CTGCATTAATGAATCGGCCAACCGCGGGGAGAGCGGTTGCGTATTGGGCGCTCTTCC 1925
QY 1739 GCTTCTCGCTCATGACTCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1798
Db 1924 GCTTCTCGCTCATGACTCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1865
QY 1799 CACTCAAAAGGCGTAAATACGGTTATCCACAGATCAGGGGATAACGCGAGGAAGAACATG 1858
Db 1864 CACTCAAAAGGCGTAAATACGGTTATCCACAGATCAGGGGATAACGCGAGGAAGAACATG 1805
QY 1859 TGACCAAAAGGCGCAAAAGGCCAGGAACCGTAAAAAGGCGCGTTGCTGCGGTTTTTC 1918
Db 1804 TGACCAAAAGGCGCAAAAGGCCAGGAACCGTAAAAAGGCGCGTTGCTGCGGTTTTTC 1745
QY 1919 CATAGGCTCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCA 1978
Db 1744 CATAGGCTCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCA 1685
QY 1979 AACCCGACGAGCTATAAAGATACAGGCTTTCCCTCGGAAGCTCCCTCGCGCTCT 2038
Db 1684 AACCCGACGAGCTATAAAGATACAGGCTTTCCCTCGGAAGCTCCCTCGCGCTCT 1625
QY 2039 CTTGTTCCGACCTCCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTGGGAAGCGTG 2098
Db 1624 CTTGTTCCGACCTCCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTGGGAAGCGTG 1565
QY 2099 GGGCTTTCTCATAGCTCAGGCTGATGATCTCAGTTGCTGCTGCTGCTGCTGCTGCTGCT 2158
Db 1564 GGGCTTTCTCATAGCTCAGGCTGATGATCTCAGTTGCTGCTGCTGCTGCTGCTGCTGCT 1505
QY 2159 CTGGGCTGTGTGACAGAACCCCGCTTACGCCGACCGCTGCGCTTATCCCGTAACTAT 2218
Db 1504 CTGGGCTGTGTGACAGAACCCCGCTTACGCCGACCGCTGCGCTTATCCCGTAACTAT 1445
QY 2219 GGTCTTGTAGTCCAAACCCGGTAAAGACACGACTTATCGCACTGGCAGAGCCACTGGTAAC 2278
Db 1444 GGTCTTGTAGTCCAAACCCGGTAAAGACACGACTTATCGCACTGGCAGAGCCACTGGTAAC 1385
QY 2279 AGGATTACAGAGCGAGTATGTAGCGGTGCTACAGAGTCTTCAAGTGTGGGCTTAC 2338
Db 1384 AGGATTACAGAGCGAGTATGTAGCGGTGCTACAGAGTCTTCAAGTGTGGGCTTAC 1325
QY 2339 TACGGCTACACTAGAAAGACAGTATTTGGTATCTCGCTCTGCTGAAGCCAGTTACTTTC 2398
Db 1324 TACGGCTACACTAGAAAGACAGTATTTGGTATCTCGCTCTGCTGAAGCCAGTTACTTTC 1265
QY 2399 GGAATAAGAGTTGTAGTCTTTGATCCGCGCAAAACCAACCGCTGCTGCTGCTGCTGCTGCT 2458
Db 1264 GGAATAAGAGTTGTAGTCTTTGATCCGCGCAAAACCAACCGCTGCTGCTGCTGCTGCTGCT 1205
QY 2459 TTTGTTTCCAGCAGCAGATTTACGCGCAGAAAAAGGATCTCAAGAGATCCTTTGATC 2518
Db 1204 TTTGTTTCCAGCAGCAGATTTACGCGCAGAAAAAGGATCTCAAGAGATCCTTTGATC 1145
QY 2519 TTTTCTACGGGTCTGACGCTCAGTGGAAACGAAAACTCACGTTAAGGGATTTTGGTCATG 2578
Db 1144 TTTTCTACGGGTCTGACGCTCAGTGGAAACGAAAACTCACGTTAAGGGATTTTGGTCATG 1085
QY 2579 AGATTATCAAAAGGATCTTCACTAGATCTTTTAAATTTAAAAATGAAGTTTAAATCA 2638
Db 1084 AGATTATCAAAAGGATCTTCACTAGATCTTTTAAATTTAAAAATGAAGTTTAAATCA 1025
QY 2639 ATCTAAGCTATATGAGTAACTTTGCTGACAGTTACCAATGCTTATCAGTGAGGCA 2698
Db 1024 ATCTAAGCTATATGAGTAACTTTGCTGACAGTTTACCAATGCTTATCAGTGAGGCA 965
|||||

```
QY 2699 CTTATCTCAGCGATCTGTCTATTGTTGTTTCATTCATAGTTGCTGACTCCCGTCGTGTAG 2758
DB 964 CTTATCTCAGCGATCTGTCTATTGTTGTTTCATTCATAGTTGCTGACTCCCGTCGTGTAG 905
QY 2759 ATAACTACGATACGGAGGCTTACCATCTGGCCCGAGTCTGCTCAATGATACCGCGAGAC 2818
DB 904 ATAACTACGATACGGAGGCTTACCATCTGGCCCGAGTCTGCTCAATGATACCGCGAGAC 845
QY 2819 CCAGCTCACCCTCTCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGCGCGAGCGC 2878
DB 844 CCAGCTCACCCTCTCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGCGCGAGCGC 785
QY 2879 AGAAGTGGTCTCAGATTTATCAGCTCCATCCAGTCTATTAATTTGTTCCCGGGAAGCT 2938
DB 784 AGAAGTGGTCTCAGATTTATCAGCTCCATCCAGTCTATTAATTTGTTCCCGGGAAGCT 725
QY 2939 AGAAGTAAAGTAGTTGCGCAGTTAATAGTTTGCAGCAAGTTGTTGCCATTTGTTACAGGCATC 2998
DB 724 AGAAGTAAAGTAGTTGCGCAGTTAATAGTTTGCAGCAAGTTGTTGCCATTTGTTACAGGCATC 665
QY 2999 GTGGTGTACGCTCGTCTGTTGATGCTTCAATCAGCTCCGCTTCCCAACGATCAAGG 3058
DB 664 GTGGTGTACGCTCGTCTGTTGATGCTTCAATCAGCTCCGCTTCCCAACGATCAAGG 605
QY 3059 CGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGTCTCTCGTCTCTCGATC 3118
DB 604 CGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGTCTCTCGTCTCTCGATC 545
QY 3119 GTTGTGAGAGTAGTTGCGCGAGTGTATCACTCATGTTATGTCAGCAGTGCATAT 3178
DB 544 GTTGTGAGAGTAGTTGCGCGAGTGTATCACTCATGTTATGTCAGCAGTGCATAT 485
QY 3179 TCTCTTACTGTCTATGCTATCGGTAAGATGCTTTTCTGTGACTCGTGTGAGTCAACCAAG 3238
DB 484 TCTCTTACTGTCTATGCTATCGGTAAGATGCTTTTCTGTGACTCGTGTGAGTCAACCAAG 425
QY 3239 TCATTCGAGAAATAGTTATGCGCGACCGAGTTGCTCTTTCGCGCGGTCATATACGGAT 3298
DB 424 TCATTCGAGAAATAGTTATGCGCGACCGAGTTGCTCTTTCGCGCGGTCATATACGGAT 365
QY 3299 ATACCGCGCCACATAGCAGAACTTTAAAGTCTCATCATTTGGAAGAGCTTCTCGGGG 3358
DB 364 ATACCGCGCCACATAGCAGAACTTTAAAGTCTCATCATTTGGAAGAGCTTCTCGGGG 305
QY 3359 CGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATTAACCCACTCGTGCA 3418
DB 304 CGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATTAACCCACTCGTGCA 245
QY 3419 CCCAACTGATCTTCAGCATCTTTTACTTTTCAACGAGTCTTCTGGGTGAGCAAAAAAGG 3478
DB 244 CCCAACTGATCTTCAGCATCTTTTACTTTTCAACGAGTCTTCTGGGTGAGCAAAAAAGG 185
QY 3479 AGCAAAATGCCCAAAAGGGAATAGGCGGACACGGAATGTTGAATACTCATCTC 3538
DB 184 AGCAAAATGCCCAAAAGGGAATAGGCGGACACGGAATGTTGAATACTCATCTC 125
QY 3539 TTTCTTTTCAATATTATTAAGCATTTTATCAGGGTTATTGTTCTCATGAGCGGATACATA 3598
DB 124 TTTCTTTTCAATATTATTAAGCATTTTATCAGGGTTATTGTTCTCATGAGCGGATACATA 65
QY 3599 TTTGAATGTTATTTAGAAAAATAAACAATAGGGGTTCCGCGCATATTTCCCGAAAAAGTG 3658
DB 64 TTTGAATGTTATTTAGAAAAATAAACAATAGGGGTTCCGCGCATATTTCCCGAAAAAGTG 5
QY 3659 CCAC 3662
DB 4 CCAC 1
```

RESULT 3
ASPGREEN1/c
LOCUS

ASPGREEN1 3626 bp DNA linear SYN 15-MAY-1998

```
DEFINITION Artificial sequences, plasmid vector pGreen-1.
ACCESSION Y09373
VERSION Y09373.1
KEYWORDS GI:1684624
          blatem gene; lacZ-PhoC fusion gene; multiple cloning site; T3
          promoter.
SOURCE synthetic construct
          synthetic construct
          other sequences; artificial sequences.
ORGANISM 1
REFERENCE
AUTHORS Thaller,M.C., Berlutti,F., Schippa,S., Selan,L. and Rossolini,G.M.
TITLE Bacterial acid phosphatase gene fusions useful as targets for
        cloning-dependent insertional inactivation
JOURNAL Biotechnol. Prog. 14 (2), 241-247 (1998)
MEDLINE 98215688
PUBMED 9548775
REFERENCE 2 (bases 1 to 3626)
AUTHORS Rossolini,G.M.
TITLE Direct Submission
JOURNAL Submitted (11-NOV-1996) G.M. Rossolini, Dip. Biologia
        Moleculare-Sez., Microbiologia, Univ. di Siena, via Laterina N.8,
        I- 53100 Siena, ITALY
COMMENT Related sequences: X64444 and X52328.
FEATURES
        Location/Qualifiers
            1..3626
                /organism="synthetic construct"
                /mol_type="other DNA"
                /db_xref="taxon:32630"
                /plasmid="pGreen-1"
                /gene="blatem"
                /gene="blatem"
                /codon_start=1
                /transl_table=11
                /protein_id="CAA70547.1"
                /db_xref="GI:1684625"
                /translation="MSIQHFRVALIPFFAFCPLPVFAHPETLVKVKDAEDQLGARVGY
                ISLDLNSGKILSPFRPFPMSTFIVLLCGAVLSRI DAGQQLGRRIHYSDNLDVE
                YSPVTEKHTDGMTVRELCSAAITMSDNTAANLLTTIGGPKELTAFLEHMGGDHVRL
                DWPELNEAI PNDERDITMPVAMATTLKLLTGLLTLASRQQLIDMEADKVAQPL
                LRSALPAGFWFIADKSGAGRGSGIIAALGPDGKPSRIIVIVITGSGQATMDERNQIA
                EIGASLIKHV"
            2139..2154
                /note="pUC M13/Reverse Sequencing primer recognition site"
            2146..2934
                /gene="lacZ-PhoC fusion"
            2146..2934
                /gene="lacZ-PhoC fusion"
                /codon_start=1
                /transl_table=11
                /protein_id="CAA70548.1"
                /db_xref="GI:1684626"
                /translation="MTMITPSAQLTLTKGNKWSSTVAALAEVLVDPAIPAGNDATTK
                POLYLYKNEQALDSLKLPPPEVSGSIQFLNDQAMYEKGRMLRNLTERGKQAOADADLA
                AGCVATASGAGFYPTDKSPELYKLLTNWLEDAGDLATRSKAEHYMRIRPFAYGT
                ETCNTKDQKKSTNGSYPSGHTSIGWATLVIAEVNPNPANDSAKLEGYQLGSGRVICG
                YHWQSDVDAAIRIVGSAAVATLSDPAFQALAKAKQEPACKSQK"
            2171..2190
                /gene="lacZ-PhoC fusion"
                /note="T3 promoter"
primer_bind
        gene
        CDS
        promoter
```

ORIGIN

Query Match 72.4%; Score 2650.6; DB 12; Length 3626;
Best Local Similarity 84.0%; Pred. No. 0;
Matches 3079; Conservative 0; Mismatches 544; Indels 42; Gaps 6;

```
QY 1 CTAATTTGTAAGCGTTAATATTGTTGTTAAATTCGCGTTAAATTTTGTAAATCAGCTC 60
DB 3626 CTAATTTGTAAGCGTTAATATTGTTGTTAAATTCGCGTTAAATTTTGTAAATCAGCTC 3567
QY 61 ATTTTAAACCAATAGCGCGAAATCGCAAAATCCCTTATAATCAAAAGAAATAGACCGA 120
DB 3566 ATTTTAAACCAATAGCGCGAAATCGCAAAATCCCTTATAATCAAAAGAAATAGACCGA 3507
```

QY 121 GATAGGTTGAGTGTGTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGAAAGTGGACTC 180
DB |||||
DB 3506 GATAGGTTGAGTGTGTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGAAAGTGGACTC 3447
QY 181 CAAAGTCAAAGGCGGAAACCGTCTATCAGGGCGATGGCCCACTACGTGGAACCAATCAC 240
DB |||||
DB 3446 CAAAGTCAAAGGCGGAAACCGTCTATCAGGGCGATGGCCCACTACGTGGAACCAATCAC 3387
QY 241 CTAATCAAGTTTGGGTCGAGGTGCGGTAAAGCACTAAATCGGAACCCCTAAAGGGAG 300
DB |||||
DB 3386 CTAATCAAGTTTGGGTCGAGGTGCGGTAAAGCACTAAATCGGAACCCCTAAAGGGAG 3327
QY 301 CCCCCGATTAGAGCTTGACGGGGAAGCCGCGCAACGTGGCGAGAGAAAGGAAGGAAGAA 360
DB |||||
DB 3326 CCCCCGATTAGAGCTTGACGGGGAAGCCGCGCAACGTGGCGAGAGAAAGGAAGGAAGAA 3267
QY 361 AGCGAAGAGAGGGCGCTAGGGCGTGGCAAGTGTAGCGGTACGCTGCGGCTAACCCAC 420
DB |||||
DB 3266 AGCGAAGAGAGGGCGCTAGGGCGTGGCAAGTGTAGCGGTACGCTGCGGCTAACCCAC 3207
QY 421 CACACCCCGCGCTTAATGCGCGCTACAGGGCGGTCCCAATTCGCCAATTCAGGCTGCG 480
DB |||||
DB 3206 CACACCCCGCGCTTAATGCGCGCTACAGGGCGGTCCCAATTCGCCAATTCAGGCTGCG 3147
QY 481 CAACTGTTGGGAAGGCGGATCGGTGCGGCGCTTTCGCTATTACGCCAGCTGGCGAAAGG 540
DB |||||
DB 3146 CAACTGTTGGGAAGGCGGATCGGTGCGGCGCTTTCGCTATTACGCCAGCTGGCGAAAGG 3087
QY 541 GGGATGTGTCGAAGGCGAATTAAGTTGGTAAAGCGCAGGCGTTCGCCAGTACGAGCTTG 600
DB |||||
DB 3086 GGGATGTGTCGAAGGCGAATTAAGTTGGTAAAGCGCAGGCGTTCGCCAGTACGAGCTTG 3027
QY 601 TAAACGACGCGCAGTACGCGCGCTCGTTTCACTACGTTTTTGAACCGCTGGAGGACGG 660
DB |||||
DB 3026 TAAACGACGCGCAGTACGCGCGCTCGTTTCACTACGTTTTTGAACCGCTGGAGGACGG 2967
QY 661 GCAGACTCCGCGTGCAAAATGTGTTTACAGCGTGATGGAGCAGATGAAGATGCTCGACAC 720
DB |||||
DB 2966 GCCCCCTCGAGTTCGACGGTATCGATGAAGCTTATTTCTGTGATTTTGTGCAAAATCC 2907
QY 721 GCTGCAGAACACGCGACTAGATAACCCCTAGAAAGATAATATATTTGTGACGTACGTTAA 780
DB |||||
DB 2906 TGTTTGGCTTCGCTAACTGCGCTGAATGCGGATCGGAATGTAAATGTCGCGACAGCG 2847
QY 781 AGATAATCATGCGTAAATTTGACGATGGATCTGTATACGACTACACTATAGGCGGAT 840
DB |||||
DB 2846 GCTGAACC-----GACAAATCCGCGCGGCATCCACATCACTCTGCGAGTATAGCCGCA 2794
QY 841 TGGGTACCGGGCCCCCTCGAGGTCGACGGTATCGATAAGCTTGATATCGAAATTCCTGCG 900
DB |||||
DB 2793 AATCACC CGGCTCTGTCCGAGCTGATAACCCGTTCCAGAAATCGCATCTGATTTGCCGG 2734
QY 901 AGCCCGGGGGATCCACTAGTTCTAGAGCGCGCGCACCGCGGTGGAGTCCAGCTTTTGT 960
DB |||||
DB 2733 GTTC----ACTTCCGCGACACTAATGCGGTGCGCGAGCGGATAGACGATGACCTGACG 2678
QY 961 TCCCTTTAGTAGGGTTAATTAGATCCCATCGGTCAATTTTACGCGAGACTATCTTTCTAG 1020
DB |||||
DB 2677 GGTAAAGATCCGTTGGTGGAGGTTTTCCTGA-TCCTTTGGTATTACAGGTTTCTGTGCGG 2619
QY 1021 GGTAAATCTAGCTGATCAGGATCATATCGTGGGTCTTTTTTCCGGCTCAGTCAATCGCC 1080
DB |||||
DB 2618 TAAACGCAACGACGGATGCGCATGTAAATGTTCTTTGGCGGAGCGGTGGCAAGTCA 2559
QY 1081 CAAAGTGGCGCTATCTGGGCATCGGGAGGAAGAGCCGCGTCCGCTTTTCCCGGAGGTTG 1140
DB |||||
DB 2558 CCGGCATCCTCAATCATATTGTCAGCACTTTATACAGCTCCGCGAGAGTCTTTTTCGGTT 2499
QY 1141 AAGCGCATGGAAGAGTTTTCGAGGATGACTGCTGCTGCATTCAGCTTGAGCGAAAC 1200
DB |||||
DB 2498 ATCGGATAGCCGAATGCCCTGAAATATGCGGTGTGCCACCCCTCGCGCCAGGTCAGCA 2439

QY 1201 GCA CGCTTTACCATGATGATTCGGGAAGGTGTGGCCATGCA---CGCTTTTAAACGGTGAAC 1257
DB |||||
DB 2438 TCTGCTGTGCGCTGTTTTCGCGCTCGGTATTGCGCAGCATACGCGCTTTCTCATACATT 2379
QY 1258 TGTTCGTTCAGGCCACCTGGGATACAGTTGTCGCGGCTTTTCCGGAACAGAGTTCCGGA 1317
DB |||||
DB 2378 GCTGATCATTTTAAAAAATCTGAATACTGCCGACTTCCGGCGGTGGCGGTAAACAGTTTCAGG 2319
QY 1318 TGTTCAGCCGGAAGCGCATCAGCAACCCGAAACAATACCGCGCACAGCCGGAACACTGCCGTG 1377
DB |||||
DB 2318 CTGTGATAGCTGTGTTCAITTTTTCAGATATAATAAT-----CCGCG 2277
QY 1378 CCGGTGTCAGATTAAATGACAGCGGTGCGGCTCGGATATTACGTACGAGGACGGGT 1437
DB |||||
DB 2276 TTGGTGGTGGCATCGTTGCTGCGGGATCGCAGGATCCACTAGTTCTAGAGCGCGCGC 2217
QY 1438 ATCCTGGCTGGATCGCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGCGGCGCGG 1497
DB |||||
DB 2216 ACCGCGTGGAGCTCCAGCTTTTGTTCCTTTAGTGAOGGTTAATT-----GCGCG 2166
QY 1498 CTTCGCGTAATCATGCTCATAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCACAAATTC 1557
DB |||||
DB 2165 CTTCGCGTAAATCATGCTCATAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCACAAATTC 2106
QY 1558 ACACAACATACAGCGCGGAAGCATAAAGTGTAAAGCTGGGTGCTAATAGTGAGCTA 1617
DB |||||
DB 2105 ACACAACATACAGCGCGGAAGCATAAAGTGTAAAGCTGGGTGCTAATAGTGAGCTA 2046
QY 1618 ACTCAATTAATTGGTTGCGCTCACTGCCGCTTTCAGTCCGGNAACCTCTCGTGCCA 1677
DB |||||
DB 2045 ACTCAATTAATTGGTTGCGCTCACTGCCGCTTTCAGTCCGGNAACCTCTCGTGCCA 1986
QY 1678 GCTGCAATTAATGAATCGGCCAACCGCGGAGAGCGGTGTCGTATTGGCGCTCTTC 1737
DB |||||
DB 1985 GCTGCAATTAATGAATCGGCCAACCGCGGAGAGCGGTGTCGTATTGGCGCTCTTC 1926
QY 1738 CGCTTCTCGCTCACTGACTCGCTCGCTCGGTGCTGCGTGGCGAGCGGTATFCAGC 1797
DB |||||
DB 1925 CGCTTCTCGCTCACTGACTCGCTCGCTCGGTGCTGCGTGGCGAGCGGTATFCAGC 1866
QY 1798 TCACCTCAAAGGCGGTAAATACGGTTATCCACAGAATCAGGGGATACGAGGAAGAAACAT 1857
DB |||||
DB 1865 TCACCTCAAAGGCGGTAAATACGGTTATCCACAGAATCAGGGGATACGAGGAAGAAACAT 1806
QY 1858 GTGACAAAAGGCCAGCAAAAAGGCCAGGAACCGTAAAAAGGCCGCTGCTGGCGTTTTT 1917
DB |||||
DB 1805 GTGACAAAAGGCCAGCAAAAAGGCCAGGAACCGTAAAAAGGCCGCTGCTGGCGTTTTT 1746
QY 1918 CCATAGGCTCCGCCCTCGCTGACGAGCATCAAAAATCGACGCTCAAGTCAAGGTGGCG 1977
DB |||||
DB 1745 CCATAGGCTCCGCCCTCGCTGACGAGCATCAAAAATCGACGCTCAAGTCAAGGTGGCG 1686
QY 1978 AAACCCGACGAGCTATAAAGATACAGCGGTTTCCCGCTGGAAGCTCCCTCGTGGCTC 2037
DB |||||
DB 1695 AAACCCGACGAGCTATAAAGATACAGCGGTTTCCCGCTGGAAGCTCCCTCGTGGCTC 1626
QY 2038 TCTGCTCCGACCCCTCGCTTACCGGATACCTGTCGCTTTCTCCCTTCGGGAAGCGT 2097
DB |||||
DB 1625 TCTGCTCCGACCCCTCGCTTACCGGATACCTGTCGCTTTCTCCCTTCGGGAAGCGT 1566
QY 2098 GCGCTTTCTCATAGCTCACTGTAGGTATCTCAGTTTCGGTGTAGGTGTTGCTGCCAA 2157
DB |||||
DB 1565 GCGCTTTCTCATAGCTCACTGTAGGTATCTCAGTTTCGGTGTAGGTGTTGCTGCCAA 1506
QY 2158 GCTGGGCTGTGTCAGCAACCCCGCTTTCAGCCGACCGCTCGCTTATCCGTTAACTA 2217
DB |||||
DB 1505 GCTGGGCTGTGTCAGCAACCCCGCTTTCAGCCGACCGCTCGCTTATCCGTTAACTA 1446
QY 2218 TCGCTTTCAGTCCAAACCCCGGTAAAGACAGACTTATCGCCACTGGCAGCAGCCACTGGTAA 2277
DB |||||
DB 1445 TCGCTTTCAGTCCAAACCCCGGTAAAGACAGACTTATCGCCACTGGCAGCAGCCACTGGTAA 1386
QY 2278 CAGGATTAAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTTGAAGTGTGGCTTAA 2337

[illegible]

Db	305	GCAGAAACTCTCAAGGATCTTACCGCTGTTGGAGATCCAGTTCGATGTAAACCCACTCTGTC	246
Qy	3418	ACCCAACTGATCTTTCAGCATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAACAGG	3477
Db	245	ACCAACTGATCTTTCAGCATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAACAGG	186
Qy	3478	AAGCAAAATGCCGCAAAAAGGAATAAGGGCGACACGGAAATGTTGAATATCTCATACT	3537
Db	185	AAGCAAAATGCCGCAAAAAGGAATAAGGGCGACACGGAAATGTTGAATATCTCATACT	126
Qy	3538	CTTCTCTTTTCAATATTATTGAGCATTTTATCAGGGTTATTGTCTCATGAGCGGATACAT	3597
Db	125	CTTCTCTTTTCAATATTATTGAGCATTTTATCAGGGTTATTGTCTCATGAGCGGATACAT	66
Qy	3598	ATTTGAATGTATTATAGAAAAATAAAACAATAGGGGTTCCCGGCACATTTTCCCGAAAAAGT	3657
Db	65	ATTGTAATGTATTATAGAAAAATAAAACAATAGGGGTTCCCGGCACATTTTCCCGAAAAAGT	6
Qy	3658	GCCAC 3662	
Db	5	GCCAC 1	
RESULT 4			
LOCUS	XXU43957		
DEFINITION	Expression vector pB10T-MRz	3624 bp	DNA
ACCESSION	U43957		linear
VERSION	U43957.1	GI:4097019	
KEYWORDS			
SOURCE	Expression vector pB10T-MRz		
ORGANISM	Expression vector pB10T-MRz		
REFERENCE	1 (bases 1 to 3624)		
AUTHORS	Ruiz,J., Wu,C.H., Ito,Y. and Wu,G.Y.		
TITLE	Submitted (27-DEC-1995) Juan Ruiz, Division of		
JOURNAL	Gastroenterology-Hepatology, University of Connecticut Health		
COMMENT	Center, Farmington, CT 06030, USA		
	In vivo expression vector for the synthesis of a ribozyme against		
	the core region of the Hepatitis B virus pregenomic RNA. It does		
	not recognise the viral core protein but the core region of the		
	viral RNA. pB10T-MRz contains a tandem of ten units of the		
	ribozyme, separated between them by the target sequence recognized		
	by the ribozyme, cloned at the XbaI-HindIII site of pBluescript SK		
	(+).		
FEATURES		Location/Qualifiers	
source	1..3624		
	/organism="Expression vector pB10T-MRz"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:45892"		
misc_feature	complement(657..1425)		
	/note="multiple cloning site"		
misc_RNA	complement(721..766)		
	/product="hammerhead ribozyme"		
misc_RNA	complement(791..836)		
	/product="hammerhead ribozyme"		
misc_RNA	complement(861..906)		
	/product="hammerhead ribozyme"		
misc_RNA	complement(931..976)		
	/product="hammerhead ribozyme"		
misc_RNA	complement(1001..1046)		
	/product="hammerhead ribozyme"		
misc_RNA	complement(1071..1116)		
	/product="hammerhead ribozyme"		
misc_RNA	complement(1141..1186)		
	/product="hammerhead ribozyme"		
misc_RNA	complement(1211..1256)		
	/product="hammerhead ribozyme"		
misc_RNA	complement(1281..1326)		
	/product="hammerhead ribozyme"		
misc_RNA	complement(1351..1396)		
	/product="hammerhead ribozyme"		

promoter	complement (1482..1604)	
gene	/note="Lac-2"	
	complement (2639..3499)	
CDS	/gene="ampR"	
	complement (2639..3499)	
	/gene="ampR"	
	/function="ampicillin resistance marker"	
	/codon_start=1	
	/transl_table=11	
	/product="beta-lactamase"	
	/protein_id="AAD10137.1"	
	/db_xref="GI:4097020"	
	/translation="MSIQHFRVALIIPFFAFLCPVFAHPETLVKVKDAEDQLGARVGY IELDLSNGKILLESFPERFPMWSFKVLLCGAVLSIRIDAGQEOQLGRIRIHYSQNDLVE YSPVTEKHLTDGMTVRELCSAAITWSDNTAANLLLTITGGPKELTAFILHMGMDSVTRL DRWEPELNEAIPNDSEDTMPVAMATLRLKLLTGLLTSLASRQQLIDHWKADKVAGPL LRSLAPAGWFIADKSGAGERSGRIIALLGPDGKFSRIIVYITIGSQATMDERNRQIIA EIGASLIKHW"	
ORIGIN		
Query Match	72.2%; Score 2644.8; DB 12; Length 3624;	
Best Local Similarity	84.3%; Pred. No. 0;	
Matches 3096; Conservative	0; Mismatches 512; Indels 64; Gaps 8;	
QY	1 CTAATTTGTAAGCGTTAATATTTTGTAAAAATTCGCGTTAAATTTTGTAAATCAGCTC 60	
DB		
QY	4 CTAATTTGTAAGCGTTAATATTTTGTAAAAATTCGCGTTAAATTTTGTAAATCAGCTC 63	
DB		
QY	61 ATTTTAAACCAATAGCGCGAAATCGGCAAAATCCCTTATAATCAAAAGAAATAGACCGA 120	
DB		
QY	64 ATTTTAAACCAATAGCGCGAAATCGGCAAAATCCCTTATAATCAAAAGAAATAGACCGA 123	
DB		
QY	121 GATAGGTTGAGTGTGTTTCCAGTTTGGAAACAGAGTCCACTATTAAAGAAAGCTGCACTC 180	
DB		
QY	124 GATAGGTTGAGTGTGTTTCCAGTTTGGAAACAGAGTCCACTATTAAAGAAAGCTGCACTC 183	
DB		
QY	181 CAACGTCAAAGGGCGAAACCGCTCTATCAGGGCGATGGCCCACTACGTGAACCACTACCC 240	
DB		
QY	184 CAACGTCAAAGGGCGAAACCGCTCTATCAGGGCGATGGCCCACTACGTGAACCACTACCC 243	
DB		
QY	241 CTAATCAAGTTTTTGGGGTTCGAGTGCCTGTAAGCACTAAATCGGAACCCCTAAAGGGAG 300	
DB		
QY	244 CTAATCAAGTTTTTGGGGTTCGAGTGCCTGTAAGCACTAAATCGGAACCCCTAAAGGGAG 303	
DB		
QY	301 CCCCAGATTAGAGCTTGACGGGGAAGCCGCGCAACGTGGCGAGAAAGGAAGGAAGAA 360	
DB		
QY	304 CCCCAGATTAGAGCTTGACGGGGAAGCCGCGCAACGTGGCGAGAAAGGAAGGAAGAA 363	
DB		
QY	361 AGCGAAAGGAGCGGCGCTAGGGCGCTGGCAAGTGTAGCGGTACGCTGCGCGTAACCAAC 420	
DB		
QY	364 AGCGAAAGGAGCGGCGCTAGGGCGCTGGCAAGTGTAGCGGTACGCTGCGCGTAACCAAC 423	
DB		
QY	421 CACACCGCGCGCTTAATGCGCGCTACAGGGCGGTCCCATTCGCCATTCAGGCTGGG 480	
DB		
QY	424 CACACCGCGCGCTTAATGCGCGCTACAGGGCGGTCCCATTCGCCATTCAGGCTGGG 483	
DB		
QY	481 CAACTGTTGGGAAGGCGAGTCGTCGGGCTCTTCGCTATTACGCCAGCTGCGGAAAGG 540	
DB		
QY	484 CAACTGTTGGGAAGGCGAGTCGTCGGGCTCTTCGCTATTACGCCAGCTGCGGAAAGG 543	
DB		
QY	541 GGGATGTGTCGAAGCGCAATTAAGTTGGGTAAACGCCAGGGTTTTCCCAAGTACACGCTTG 600	
DB		
QY	544 GGGATGTGTCGAAGCGCAATTAAGTTGGGTAAACGCCAGGGTTTTCCCAAGTACACGCTTG 603	
DB		
QY	601 TAAACGACGGCCGATG-----AGGGCGCTCGTTTCAATTCAGTTTGTGAACCGGTGGA 654	
DB		
QY	604 TAAACGACGGCCGATGTTAATACGACTCACTATATAGGGCGAATTTGGGTACCGGGCC 663	
DB		
QY	655 GGACGGGACAGCTCGCGGTGCAAAATGTGTTTACAGCGGTGATGAGAGAGATCAAGATGCT 714	
DB		
QY	664 CCCCCTCGAGGTGCGAGGTATCGATAAGCTTGAGACTAGTTTACGAGAGACTCTGGATCCAA 723	
DB		
QY	715 CGACACGCTGCAGAACACGACAGCTAGATTAAACCCCTAGAAAGATAATCATATTGTGACGTA 774	
DB		

DB	724 CCTTAGAGTTTCGTCTCAGCGACTCATCAGTCCT---GAGCATTTGTCTAGTTTCAGGAGA 780	
QY	775 CGTTAAAGATATCATCGTGAATAAATTGACCGATGGATCTGTAAATACGACTCATTATAGG 834	
DB	781 CTCTGGATCCAAGCCTTAGAGTTTCGTCTCAGGACTCATCAGTCTCTGAGCAATTTGTCTA 840	
QY	835 GCGAATTGGGTACCGGGCCCCCTCGAGGTGCGAGGTATCGATAGAGTTCGATATCGAAT 894	
DB	841 GTTCAGGAGACTCTCGATCCCAAGCCTTAGAGTTTCGTCTCAGCGA-----886	
QY	895 TCCTCGAGCCCGGGGATCCACTAGTTCTTAGAGCGCGCGCCACCGGGTGGAGTCCAGC 954	
DB	887 CTCATCAGTCTCTGAGCAATTTGTCTAGTTTCAGGAGA-----C 921	
QY	955 TTTTGTTCCTTTAGTAGAGGTTAATTAGATCCCATCGTCAATTTTACGACAGATATCT 1014	
DB	922 TCTGATCAAGCCTTAGAGTTTCGTCTCAGGACTCATCAGTCTCTGAGCAATTTGTCTAG 981	
QY	1015 TTCTAGGGTTAATCTAGCTGCATCAGGATCATATCTCGGGTCTTTTTCGGCTCAGTC 1074	
DB	982 TTCAGGAG-----ACTCTGGATCCAAGCCTTAGAGTTTCGTC--CTCAGGACTCATCA 1033	
QY	1075 ATCGCCCCAAGCTGGCGCTATCTGGGCATCGGGGAGAAAGACCCGTCCTTTTCCCGG 1134	
DB	1034 GTCCTGAGCAATTTCTAGTTTCAAGGAGACTCTGGATCCAAGCCTTAGAGTTTCGTCCTCAC 1093	
QY	1135 AGTTTGAAGCGCATGGAAGAGTTTTCGGAGGATGCTGCTGCTGCAATGAGCTTGAGC 1194	
DB	1094 GGACTCATCAGTCTGAGCAATTTCTAGTTTCAAGGAGACT-CTGGATCCAAGCCTTAGAGT 1152	
QY	1195 GAAAAACGACGTTTACCATGATGATTTCGGGAAGTGTGGCCATGACGCGCTTTAAACGGTG 1254	
DB	1153 TTCGTCCTCAAGGACTCATCAGTCTCTGAGCAATTTGTCTAGTTTCAAGGAGACTCTGGATCCAA 1212	
QY	1255 AACTGTTTCGTTAGGCCACCTGGGATAACGTTTCGTCGGGCTTTTTCGGGACACAGTTTCC 1314	
DB	1213 GCCTTAGAGTTTCGTCCTCAGGACTCATCAGTCTGAGCAATTTGTCTAGTTTCAAGGAGACT 1272	
QY	1315 GGATGCTCAGCGCGGAGCGCATCAGAACCCGCAACATACCGGCGACAGCCGGAAGTCC 1374	
DB	1273 CTGGATCCAAGCCTTAGAGTTTCGTCCTCAGCGACTCATCAGTCTCTGAGCAATTTGTCTAGT 1332	
QY	1375 GTGCCGCTGTGCAGATTAAATGACAGCGGTGGCGGCTGGGATATTACGTCAGCGAGGAGC 1434	
DB	1333 TCAGGAGACTCTGGATCCAAGCCTTAGAGTTTCGTCCTCAGGACTCATCAGTCTCTGAGC 1392	
QY	1435 GGTATCTGGCTGGATCGCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGCGGG- 1493	
DB	1393 ATTGTCTAGAGCGGCGCGCACCGCGGTGGAGCTCCAGCTTTTGTTCCTTTAGTGAGGGT 1452	
QY	1494 -----CGCGCTTGGCGTAATCATGTCATAGCTGTTTCTGTGTGAATTTGTATTCGCG 1547	
DB	1453 TAAATTCAGCTTGGCGTAATCATGTCATAGCTGTTTCTGTGTGAATTTGTATTCGCG 1512	
QY	1548 TCACAATTCACACAACTACGAGCGGAGCATAAAGTGTAAAGCCTGGGGTGCTTAAT 1607	
DB	1513 TCACAATTCACACAACTACGAGCGGAGCATAAAGTGTAAAGCCTGGGGTGCTTAAT 1572	
QY	1608 GAGTGAGCTAACTCATATTAAATTGGGTTCGCTCATCTGCCGCTTTTCAAGTCGGGAACCC 1667	
DB	1573 GAGTGAGCTAACTCATATTAAATTGGGTTCGCTCATCTGCCGCTTTTCAAGTCGGGAACCC 1632	
QY	1668 TGTCTGTCCAGCTGATTAATGMAATCGGCAACCGCGGGGAGAGCGGTTTTCGCTATTG 1727	
DB	1633 TGTCTGTCCAGCTGATTAATGMAATCGGCAACCGCGGGGAGAGCGGTTTTCGCTATTG 1692	
QY	1728 GCGCTCTTTCGCTCTCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCG 1787	
DB	1693 GCGCTCTTTCGCTCTCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCG 1752	
QY	1788 CGGTATCAGCTCACTCAAAAGCGGTAATACGGTTATCCACAGAAATCAGGGGATACGCGAG 1847	
DB		

Db 1753 CCGTATCAGCTCACTCAAAGCGGTAAATACGGTTATCCACAGAAATCAGGGGATAACGCAG 1812
Qy 1848 GAAAGAACATGTGAGCAAAAGGCGCAGCAAAAGGCCAGGAACCGTAAAAAGGCGCGTTGC 1907
Db 1813 GAAAGAACATGTGAGCAAAAGGCGCAGCAAAAGGCCAGGAACCGTAAAAAGGCGCGTTGC 1872
Qy 1908 TGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCAGCGTCAAGTC 1967
Db 1873 TGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCAGCGTCAAGTC 1932
Qy 1968 AGAGGTGGCGAAACCCGACGAGTATAAAGATACCAAGCGTTTCCCGCTGGAAGCTCCC 2027
Db 1933 AGAGGTGGCGAAACCCGACGAGTATAAAGATACCAAGCGTTTCCCGCTGGAAGCTCCC 1992
Qy 2028 TCGTGGCTCTCTGTGTTCCGACCCCTGCTTACCGGATACCTGTCGCGCTTTCTCCCTT 2087
Db 1993 TCGTGGCTCTCTGTGTTCCGACCCCTGCTTACCGGATACCTGTCGCGCTTTCTCCCTT 2052
Qy 2088 CGGGAAGCGTGGCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCCGTTGAGGTGC 2147
Db 2053 CGGGAAGCGTGGCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCCGTTGAGGTGC 2112
Qy 2148 TTCCGCTCCAAGCTGGGCTGTGTGACGAAACCCCGCTTCCAGCCGACCGCTGCGCTTAT 2207
Db 2113 TTCCGCTCCAAGCTGGGCTGTGTGACGAAACCCCGCTTCCAGCCGACCGCTGCGCTTAT 2172
Qy 2208 CCGGTAACCTATCGTCTTTGAGTCAACCCCGTTAAGACACGACTTATCGCACTGGCAGCAG 2267
Db 2173 CCGGTAACCTATCGTCTTTGAGTCAACCCCGTTAAGACACGACTTATCGCACTGGCAGCAG 2232
Qy 2268 CCACTGGTAACAGGATTACGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGT 2327
Db 2233 CCACTGGTAACAGGATTACGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGT 2292
Qy 2328 GGTGGCCTAACTACGCTACCTAGAGGACAGTATTTGCTATCTCGCTCTGCTGAAGC 2387
Db 2293 GGTGGCCTAACTACGCTACCTAGAGGACAGTATTTGCTATCTCGCTCTGCTGAAGC 2352
Qy 2388 CAGTTACCTTCGGAAAGAGTTGGTAGCTCTTGATCCGGCAAAACCAACCGCTGGTA 2447
Db 2353 CAGTTACCTTCGGAAAGAGTTGGTAGCTCTTGATCCGGCAAAACCAACCGCTGGTA 2412
Qy 2448 GCGTGGTTTTTTGTTTGAAGCAGCAGATTAACGGCGAGAAAAGAGATCTCAAGAG 2507
Db 2413 GCGTGGTTTTTTGTTTGAAGCAGCAGATTAACGGCGAGAAAAGAGATCTCAAGAG 2472
Qy 2508 ATCCCTTTGATCTTTCTACGGGCTCTGACGCTCAGTGGAAACGAAACTCACGTTAAGGA 2567
Db 2473 ATCCCTTTGATCTTTCTACGGGCTCTGACGCTCAGTGGAAACGAAACTCACGTTAAGGA 2532
Qy 2568 TTTTGGTCTAGATTAATCAAAAGGATCTTCACTAGATCTCTTTTAAATTTAAATGAA 2627
Db 2533 TTTTGGTCTAGATTAATCAAAAGGATCTTCACTAGATCTCTTTTAAATTTAAATGAA 2592
Qy 2628 GTTTTAAATCAATCTAAAGTATATATGAGTAACTTGGTCTGACAGTTACCAATGCTTAA 2687
Db 2593 GTTTTAAATCAATCTAAAGTATATATGAGTAACTTGGTCTGACAGTTACCAATGCTTAA 2652
Qy 2688 TCAGTGAGGACCTATCTCAGCGATCTGTCTATTTCTGTTTCATCATAGTTGCTGACTCC 2747
Db 2653 TCAGTGAGGACCTATCTCAGCGATCTGTCTATTTCTGTTTCATCATAGTTGCTGACTCC 2712
Qy 2748 CCGTCTGTAGATAAATACGATACGGAGGCTTACCATCTGGCCCCAGTGTGCAATGA 2807
Db 2713 CCGTCTGTAGATAAATACGATACGGAGGCTTACCATCTGGCCCCAGTGTGCAATGA 2772
Qy 2808 TACCGGAGACCCAGCTCAGGCTCCAGATTTATCAGCAATAAACCAGCGACCGGAA 2867
Db 2773 TACCGGAGACCCAGCTCAGGCTCCAGATTTATCAGCAATAAACCAGCGACCGGAA 2832
Qy 2868 GGGCCGAGCGCAGAGTGGTCTGCAACTTATCCGCTCCATCCAGTCTATTAATGTT 2927
Db 2833 GGGCCGAGCGCAGAGTGGTCTGCAACTTATCCGCTCCATCCAGTCTATTAATGTT 2892

Qy 2928 GCCGGAAGCTAGAGTAAGTAGTTCCCGAGTTAAATAGTTTGGCAACAGTTGTTGCCATTG 2987
Db 2893 GCCGGAAGCTAGAGTAAGTAGTTCCCGAGTTAAATAGTTTGGCAACAGTTGTTGCCATTG 2952
Qy 2988 CTACAGGCATCTGGTGTCCAGCTCGTCTTGGTATGCTTCACTTCACTCGGTTCC 3047
Db 2953 CTACAGGCATCTGGTGTCCAGCTCGTCTTGGTATGCTTCACTTCACTCGGTTCC 3012
Qy 3048 AACGATCAAGCGAGTTACATGATCCCATCTGTGTGCAAAAAGCGTTAGCTCTTCG 3107
Db 3013 AACGATCAAGCGAGTTACATGATCCCATCTGTGTGCAAAAAGCGTTAGCTCTTCG 3072
Qy 3108 GTCTCTCCGATCTGTGTCAGAGTAAGTTGGCCGAGTGTATCATTGATGTTATGGCAG 3167
Db 3073 GTCTCTCCGATCTGTGTCAGAGTAAGTTGGCCGAGTGTATCATTGATGTTATGGCAG 3132
Qy 3168 CACTGCATAATCTCTTACTGTGTCATGCCATCCGTAAGATGCTTTCTGTGACTGGTGAGT 3227
Db 3133 CACTGCATAATCTCTTACTGTGTCATGCCATCCGTAAGATGCTTTCTGTGACTGGTGAGT 3192
Qy 3228 ACTCAACCAAGTCATTTCTGAGAAATAGTATGCGGCGACCGAGTTGCTCTTGCCTGGCGT 3287
Db 3193 ACTCAACCAAGTCATTTCTGAGAAATAGTATGCGGCGACCGAGTTGCTCTTGCCTGGCGT 3252
Qy 3288 CAATACGGGATTAATACCGGCGCACATAGCAGAACTTTTAAAGTGTCTCATCTTGGAAAC 3347
Db 3253 CAATACGGGATTAATACCGGCGCACATAGCAGAACTTTTAAAGTGTCTCATCTTGGAAAC 3312
Qy 3348 GTTCTTCGGGGGCAAACTCTCAAGGATCTTACCGCTGTGAGATCCAGTTCGATGTAAC 3407
Db 3313 GTTCTTCGGGGGCAAACTCTCAAGGATCTTACCGCTGTGAGATCCAGTTCGATGTAAC 3372
Qy 3408 CCACTCGTCACCCCAACTGATCTTCAGCATCTTTTACTTTTACCAGCGTTTCTGGGTGAG 3467
Db 3373 CCACTCGTCACCCCAACTGATCTTCAGCATCTTTTACTTTTACCAGCGTTTCTGGGTGAG 3432
Qy 3468 CAAAAACAGGAAGCAAAATGCCCAAAAGGAATAAGGGCGACACGGAATGTTGAA 3527
Db 3433 CAAAAACAGGAAGCAAAATGCCCAAAAGGAATAAGGGCGACACGGAATGTTGAA 3492
Qy 3528 TACTCATCTCTCTCTTTTCAATATTATTGAAGCATTTATCAGGTTTATGTTCTCATGA 3587
Db 3493 TACTCATCTCTCTCTTTTCAATATTATTGAAGCATTTATCAGGTTTATGTTCTCATGA 3552
Qy 3588 GCGGATACATATTTGAATGTTATTAGAAAATAAACAATAGGGTTCCGCGCATTTTC 3647
Db 3553 GCGGATACATATTTGAATGTTATTAGAAAATAAACAATAGGGTTCCGCGCATTTTC 3612
Qy 3648 CCCGAAAAGTGC 3659
Db 3613 CCCGAAAAGTGC 3624

RESULT 5

A20702 LOCUS 3681 bp DNA linear PAT 22-AUG-1994
DEFINITION pKSEL5 DNA sequence.
ACCESSION A20702
VERSION A20702.1 GI:583404
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 3681)
AUTHORS
TITLE CARRIER-BOUND RECOMBINANT PROTEIN, PROCESS FOR PRODUCING IT AND ITS
JOURNAL USE AS AN IMMUNOGEN AND VACCINE
Patent: WO 9113155-A 6 05-SEP-1991;
FEATURES
Location/Qualifiers
source
1..3681
/organism="synthetic construct"
/mol_type="unassigned DNA"

ORIGIN		/db_xref="taxon:32630"	
Query Match		72.1%; Score 2639; DB 6; Length 3681;	
Best Local Similarity		83.5%; Pred. No. 0;	
Matches 3075; Conservative		0; Mismatches 580; Indels 28; Gaps 6;	
QY	3	AAATTGTAGCGTTAAATATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTCAT	62
DB	1	AAATTGTAAAGTTAAATATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTCAT	60
QY	63	TTTTTAACCAATAGCGCGAAATTCGCAAAATCCCTTATATAATCAAAAGAAATAGACCGAGA	122
DB	61	TTTTTAACCAATAGCGCGAAATTCGCAAAATCCCTTATATAATCAAAAGAAATAGACCGAGA	120
QY	123	TAGGTTGAGTGTGTTCCAGTTTGGAAACAAAGTCCACTATTAAGAACGTGGACTCCA	182
DB	121	TAGGTTGAGTGTGTTCCAGTTTGGAAACAAAGTCCACTATTAAGAACGTGGACTCCA	180
QY	183	ACGTCAAAAGGCGAAAAACCGTCTATCAGGCGATGGCCCACTACGTGAACCATCACCT	242
DB	181	ACGTCAAAAGGCGAAAAACCGTCTATCAGGCGATGGCCCACTACGTGAACCATCACCT	240
QY	243	AATCAAGTTTTTGGGTCAGGTCGCGTAAAGCACTAAATCGAAACCTTAAAGGAGCC	302
DB	241	AATCAAGTTTTTGGGTCAGGTCGCGTAAAGCACTAAATCGAAACCTTAAAGGAGCC	300
QY	303	CCGATTTAGACTTACGGGGAAGACCGCGAAACGTGGCGAAGAGGAGGAAGAAG	362
DB	301	CCGATTTAGACTTACGGGGAAGACCGCGAAACGTGGCGAAGAGGAGGAAGAAG	360
QY	363	CGAAAGGAGCGCGCTAGGCGCTGGCAAGTGTAGCGGTACAGCTGCGCGTAAACACCA	422
DB	361	CGAAAGGAGCGCGCTAGGCGCTGGCAAGTGTAGCGGTACAGCTGCGCGTAAACACCA	420
QY	423	CACCGCGCGCTTAATGCGCGCTACAGGGCGCGTCCCATTTGCCATTCAGGCTGCBA	482
DB	421	CACCGCGCGCTTAATGCGCGCTACAGGGCGCGTCCCATTTGCCATTCAGGCTGCBA	480
QY	483	ACTGTTGGAGGCGGATCGGTCGGGCTCTTCGCTATTACGCCAGCTGGCGAAGGGG	542
DB	481	ACTGTTGGAGGCGGATCGGTCGGGCTCTTCGCTATTACGCCAGCTGGCGAAGGGG	540
QY	543	GATGTGCTCAAGGCGATTAAAGTTGGGTAACCCACAGGGTTTTCCAGTCAACAGCTGTGA	602
DB	541	GATGTGCTCAAGGCGATTAAAGTTGGGTAACCCACAGGGTTTTCCAGTCAACAGCTGTGA	600
QY	603	AAACGAGCCAGTGAAGCGCGCTCGTTCAATTCACGTTTTTGAACCCCGTGGAGGACGGG	662
DB	601	AAACGAGCCAGTGAATTTGTAATACGACTCACTATAGGGCGAATTTGGAGCTCCACCGCG	660
QY	663	AGACTCGCGGTCAAAATGTGTTTACAGGCGATGGAGGAGATGAAGATGCTCGACACGC	722
DB	661	GTGGCGCGCGCTAGTAGTGGTGCATCTCAGTACAATCTGCTCTGATGCCGCATAGTTA	720
QY	723	TGCAGAACACGAGCTAGATTAAACCTTAGAAGATAATCATATTGTGACGTACGTTAAAG	782
DB	721	AGCCAGTATATACATCCGCTATCGTACGTACGTGGGTATGGCTGGCCCGACACCC	780
QY	783	ATAATCATCGCTPAAATTTGACGATGGGATCTGTAATACGACTCACTATAGGGCGAAATG	842
DB	781	GCCAAACACCGCTGACGCGCCCTGACGCGCTGTCTGCTCCGGCATCCGCTTACAGACA	840
QY	843	GGTACCGGGCCCCCTCGAGTCCAGCGATTCGA-----TAAGCTTGATATCAATTC	896
DB	841	AGCTGTGACCGCTCCCGGAGCTGATGTGTAGAGGTTTTTCCCGTCAATCACCGAAACG	900
QY	897	CTGCAGCCCGGGGATCCACTAGTTCTAGAGCGGCGCCACCGCTGGAGCTCCAGCTT	956
DB	901	CCGAGGCGAGTAAGGTCCGATCTTTGTGAGCAATTCGTCCTTAAAGTAAGCAATTGCTG	960
QY	957	TTGTTCCCTTTAGTGAGGGTTAATTAGATGCCAT-----GCGTCAATTTTACGAGAC	1009
DB			

961	TAAAGTCGTCAC	TGTGCGGATCACCGCTTCCAGTAGCGACAGAGCAATTTGATTGGTAAA	1020
1010	TATCTTTCTAGGGTAA	TCTAGCTCATCAGGATCATATCGTCGGTCTTTTTTCCGGCT	1069
1021	TTTCAGAGAAAGAT	TCGGAGGAAGATCAATACATAAAGAGTTGAACCTTCTTTGTGTCT	1080
1070	CAGTCATCGCCCAAGCT	TGCGCTATCTCGGGCATCGGGGAGGAAGCCGCTGCGCTTTTC	1129
1081	TCGACATGGTAA	TCTCATGTTTGAATGGCCCTAGAGGATCCCGCCCAAGCTTGCATGCC	1140
1130	CGCGAGGTTGA	CGCGCATGGAAGAGTTTCCCGAGGATGACTCTGCTGCTGATTTGACGT	1189
1141	TGCGGTACCAT	CTTAGAGGATCCCGCAGCTCGACGCCATTAATAATGTTTTCGCTAAAT	1200
1190	TCAGCGAAAAC	CGCAGTTTACCATGATTTCCGGAAGSTGTGGCATGCACGCTTTAA	1249
1201	TCAGCGCTT	CCATGATGAGACAGGCCGTTT--GAATGTTGACGGGATGAACATAATAG	1258
1250	CGGTGAAC	CTGTTCTCAGGGCACCTGGGATACCAAGTTCTGTCGCGGCTTTTCCGGACACA	1309
1259	CAATGACGCGACGA	ATAAATCAACAGG--AGCAGGAAGCGAGGGTATCCCAACAAGT	1315
1310	GTTCGGA	TGFTACGCCCGAGCGCATCAGCAACCCGAAACAATACCGCGGACAG-----	1363
1316	CCAGCGTACCA	TAAACGCAAGCCTCAACGCGACGACGAGAGCGGTCACTAGTACAA	1375
1364	CGGAAC	CTGCCGCTGCGGTGTCAGATTAAATGACAGCGGTGCGCGCTGGGATATTACGT	1423
1376	TCCAAAC	CTTCTGTCAGAAATCGAAATCATCTTCGGTTAAATCCAAACCGGAC	1435
1424	CAGCGAGGAC	CGGTAT----CTGCTGGATGCCGAGAAATGGACATGGATACCCCGTG	1479
1436	AAGCTGAAT	GAGAATTCGACCTCGAGGGGGCCCGGTACCCAGCTTTTGTTCCTTTA	1495
1480	AGTTACCGCGG	CGCGCTTGGCGTAAATCATGCTAGTGTCTGTTCTGCTGTAATG	1539
1496	GTGAGGGT	TAAATTCGAGCTTGGCGTAAATCATGCTCATGCTTCTCTGTGTAATG	1555
1540	TTATCCGCT	CAAAATTCACACAACATACGAGCGGAGGATATAAGTGTAAAGCTGGG	1599
1556	TTATCCGCT	CAAAATTCACACAACATAGGAGCGGAAAGCATAAAGTGTAAAGCTGGG	1615
1600	TGCTTAAT	GAGTGAAGTAACTCACAATAATTCGCTTGGCTCACTGCCCGCTTTCCAGTC	1659
1616	TGCTTAAT	GAGTGAAGTAACTCACAATAATTCGCTTGGCTCACTGCCCGCTTTCCAGTC	1675
1660	GGGAAAC	CTGCTGCGCAGCTGATTAATGAAATCGGCAACGCGCGGGGAGGCGGTTT	1719
1676	GGGAAAC	CTGCTGCGCAGCTGATTAATGAAATCGGCAACGCGCGGGGAGGCGGTTT	1735
1720	CGGTAT	TGCGCGCTTCTCGCTCACTGACCTCGCTCGCTCGCTCGCTCGCTCGCT	1779
1736	CGGTAT	TGCGCGCTTCTCGCTCACTGACCTCGCTCGCTCGCTCGCTCGCTCGCT	1795
1780	CGCGCAGCGG	TATCAGCTCACTCAAAGCGGTAAATCGGTTATCCACAGATACAGGGA	1839
1796	CGCGCAGCGG	TATCAGCTCACTCAAAGCGGTAAATCGGTTATCCACAGATACAGGGA	1855
1840	TAAACGAGAA	AGAAATGAGCAAAAGGCGCAAAAGGCGGAGGAGGAGGAGGAGGAGGAGG	1899
1856	TAAACGAGAA	AGAAATGAGCAAAAGGCGCAAAAGGCGGAGGAGGAGGAGGAGGAGGAGG	1915
1900	CGCGT	TGCTGCGTTTTTCCATAGGCTCGCCCTCGACGAGCATCAAAATTCGACG	1959
1916	CGCGT	TGCTGCGTTTTTCCATAGGCTCGCCCTCGACGAGCATCAAAATTCGACG	1975
1960	CTCAAGT	CAGAGTGGCGAAACCCGACAGGACTATAAAGATAACAGGCGTTTCCCGCTG	2019
1976	CTCAAGT	CAGAGTGGCGAAACCCGACAGGACTATAAAGATAACAGGCGTTTCCCGCTG	2035
2020	AAGCT	CTCTGCTGCTTCTGTCGACCGCTGCTTACCGATACCTGCTGCGCTT	2079
2036	AAGCT	CTCTGCTGCTTCTGTCGACCGCTTACCGATACCTGCTGCGCTT	2095

183 QY ACCTCAAAGGCGAAAAACCGTCTATCAGGCGATGGCCCACTACGTGAACCATCACCCCT 242
181 Db AGTCAAAGGCGAAAAACCGTCTATCAGGCGATGGCCCACTACGTGAACCATCACCCCT 240
243 QY AATCAAGTTTTTTGGGGTTCAGGTCGCGTAAAGCACTAAATCGGAACCCCTAAAGGAGCC 302
241 Db AATCAAGTTTTTTGGGGTTCAGGTCGCGTAAAGCACTAAATCGGAACCCCTAAAGGAGCC 300
303 QY CCGGATTTAGCTTGAACGGGAAAGCCGGCGAAGCTGGCGAAGAGGAGGAGAAAG 362
301 Db CCGGATTTAGCTTGAACGGGAAAGCCGGCGAAGCTGGCGAAGAGGAGGAGAAAG 360
363 QY CGAAAGGAGCGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACTGCGCGTAAACCA 422
361 Db CGAAAGGAGCGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACTGCGCGTAAACCA 420
423 QY CACCGCGCGCTTAAATGCGCGCTACAGGGCGCGTCCCATTCGCAATTCAGGCTGCGCA 482
421 Db CACCGCGCGCTTAAATGCGCGCTACAGGGCGCGTCCCATTCGCAATTCAGGCTACGCA 480
483 QY ACTGTGGGAAGGCGATCGGTCGGGCTCTTCGCTATTACGCCAGCTGGCGAAGGGG 542
481 Db ACTGTGGGAAGGCGATCGGTCGGGCTCTTCGCTATTACGCCAGCTGGCGAAGGGG 540
543 QY GATGTGCTGCAAGGCGATTTAAGTTGGGTAAACCGCAGGGTTTTCCAGTCAACGACGTTGTA 602
541 Db GATGTGCTGCAAGGCGATTTAAGTTGGGTAAACCGCAGGGTTTTCCAGTCAACGACGTTGTA 600
603 QY AAACGACGCGCATGAGCGCGCTCGTTCAITTCAGCGTTTTGAAACCGGTGGAGGACGGG 662
601 Db AAACGACGCGCATGAGCGCGCTCGTTCAITTCAGCGTTTTGAAACCGGTGGAGGACGGG 660
663 QY AGACTCGCGGTCAAAATGTGTTTTACAGCGTATGGAGCAGATGAAGATGTCGACACGC 722
661 Db GTGGCGGCGCTCTAGTATGGTGCATCTCAGTACAAATCTGCTCTGATGCGCATAGTTA 720
723 QY TGCAGAAACGCGAGCTAGATTAAACCTTAGAAGATAAATCATATTGTGACGTACGTTAAAG 782
721 Db AGCCAGTATATACATCCGCTATCGTACGTACGTTGGTTCATGGCTGCGCCCGACACCC 780
783 QY ATAATCATGCGTAAATTTGACGATCGGATCTGTAATACGATCACTATAGGGCGAAATTG 842
781 Db GCCAACACCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGCATCCGCTTACAGACA 840
843 QY GTACCGGGCCCCCTCGAGGTCGACGATCGA- ----TAAGCTTGATATCGAATTC 896
841 Db AGCTGTGACCGTCTCCGGAGCTGCATGTGTGAGAGGTTTTACCGTCAATCACCGAAAGC 900
897 QY CTGCAGCCCGGGGATCCCACTAGTTCTAGAGCGGCGCCACCGCGTGGAGCTCCAGCTT 956
901 Db CGCGAGGCGAGTAAGGTCGATGCTTTGTGAGCAATTCGTCCTTAAAGTAAGCAATTGCTG 960
957 QY TTGTTCCCTTTAGTGAGGGTTAAATTAGATCCCAT- ----GCGTCAATTTTACGAGAC 1009
961 Db TAAAGTCGTCACTGTGCGGATCACCGCTTCCAGTAGCGACAGAGAATTCATTGGTAAA 1020
1010 QY TATCTTTCTAGGGTTAATCTAGCTGATCAGGATCATATCGTCGGGTCTTTTTCGGCT 1069
1021 Db TTTCAGAGAAAGATCGCGAGGAAGATCAATACATAAAGAGTTGAATCTTTTGTGTCT 1080
1070 QY CAGTCATCGCCCAAGCTGCGCTATCTGGGCATCGGGAGGAAGAGCCGCTGCTTTTC 1129
1081 Db TCGACATGGGTAACTCTCATGTTGTAATGGCCCTAGAGGATCCGGCCCAAGCTTCATGCC 1140
1130 QY CGCGAGGTTGAAGCGGATGAAAGAGTTTCCGAGGATGATCTGCTGTCGATGACGT 1189
1141 Db TCGAGGTGACCTCTAGAGGATCCCGGACGCTCGACGCCAATTAATAATGTTTTCCGTAAT 1200
1190 QY TGAGCGAAAAACGACGTTTTACCATGATGATTCGGGAAGGTGCGCATGCGACGCTTTAA 1249
1201 Db TCAGCGCCTTTCATGATGAGACAGGCGCGTTT- --GAATGTTGAGGGAGTAACATTAATAG 1258

1250 QY CGGTGAACCTGTTGTTTCCAGGCCACCTGGGATACCAAGTTCTGTCGCGGCTTTTCCGACACA 1309
1259 Db CAATGACGCGCAGCAATAAATCAACAGG- --AGCAGGAAAGCGAGGTTATCCCAAAAGT 1315
1310 QY GTTCCGGAATGTCAGCCCGAAGCGCATCAGCAACCCGAAACAAATACCGGGGACAG- ---- 1363
1316 Db CCAGCGTACCAATAAACGCAAGCCTCAACGCGACGACGAGACGCGTCAAGTAGCA 1375
1364 QY CCGGAACCTGCGGTGCGGTGTGAGATTAAATACAGAGCGGTGCGGCTGGGATATTACGT 1423
1376 Db TCCAAACTTTGTTACTCGTCAGAAATCGAAATCATCTTCGTTAAATCCAAAAACGCGAG 1435
1424 QY CAGCAGGACCGGGTAT- ----CCTGCTGATGCCAGCAAGAAATGGACATACGATACCCCGTG 1479
1436 Db AAGCCTGAATGAGAAATTCGACCTCGAGGGGGCGCGGTACCCAGCTTTTGTTCCTTTTA 1495
1480 QY AGTTACCGCGGGGCGCTTGGCGTAAATCATGTGTCATAGCTTTCTCTGTGTGAATTTG 1539
1496 Db GTAGGGTTAAATTCGAGCTTGGCGTAAATCATGTGTCATAGCTTTTCTCTGTGTGAATTTG 1555
1540 QY TTATCCGCTCAAAATTCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGG 1599
1556 Db TTATCCGCTCAAAATTCACACAACATAGGAGCCGGAAGCATAAAGTGTAAAGCCTGGGG 1615
1600 QY TGCCTAAATGAGTGAAGCTAACTCAATTAATTGCTGCTCACTGCCCGCTTTCCAGTTC 1659
1616 Db TGCCTAAATGAGTGAAGTAACTCAATTAATTGCTGCTCACTGCCCGCTTTCCAGTTC 1675
1660 QY GGGAAACCTGCTGTCGAGCTGCATTAAATGAATCGGCCAAACGCGCGGGAGAGCGGTTT 1719
1676 Db GGGAAACCTGCTGTCGAGCTGCATTAAATGAATCGGCCAAACGCGCGGGAGAGCGGTTT 1735
1720 QY GCGTATTCGGGCGCTCTTCGCTTCTCGCTCACTCACTCGCTCGCTCGGTCGTTCCGCT 1779
1736 Db GCGTATTCGGGCGCTCTTCGCTTCTCGCTCACTCACTCGCTCGCTCGGTCGTTCCGCT 1795
1780 QY GCGGAGCGGTATCAGCTCACTCAAAGCGGTAAATCGGTTATCCACAGAAATCAGGGA 1839
1796 Db GCGGAGCGGTATCAGCTCACTCAAAGCGGTAAATCGGTTATCCACAGAAATCAGGGA 1855
1840 QY TAAACGAGAAAGAACATGTAGCAAAAGGCGCAAAAGGCCAGGAAACCGTTAAAGAGC 1899
1856 Db TAAACGAGAAAGAACATGTAGCAAAAGGCGCAAAAGGCCAGGAAACCGTTAAAGAGC 1915
1900 QY CGCGTTGCTGCGTTTTTCCATAGGCTCGGCCCCCTGACGAGCATCAAAAAATTCGACG 1959
1916 Db CGCGTTGCTGCGTTTTTCCATAGGCTCGGCCCCCTGACGAGCATCAAAAAATTCGACG 1975
1960 QY CTCAGTCAGAGGTGGGAAACCCGACAGGACTATAAAGATACAGGGGTTTTCCCGCTGG 2019
1976 Db CTCAGTCAGAGGTGGGAAACCCGACAGGACTATAAAGATACAGGGGTTTTCCCGCTGG 2035
2020 QY AAGTCCCTCGTGGCTCTCTGTTCCGACCTCGCGCTTACCGGATACCTGTCGCGCTT 2079
2036 Db AAGTCCCTCGTGGCTCTCTGTTCCGACCTCGCGCTTACCGGATACCTGTCGCGCTT 2095
2080 QY TCTCCCTTCGGAAGCGTGGCGCTTTCTCATAGCTCAAGCTGTAGGTATCTCAGTTCCGT 2139
2096 Db TCTCCCTTCGGAAGCGTGGCGCTTTCTCATGTCTACGCTGTAGGTATCTCAGTTCCGT 2155
2140 QY GTAGTCGTTTCGCTCAAAGCTGGGCTGTGTGCAAGAACCCCGCTTCAGCCGACGCGTG 2199
2156 Db GTAGTCGTTTCGCTCAAAGCTGGGCTGTGTGCAAGAACCCCGCTTCAGCCGACGCGTG 2215
2200 QY CGCCTTATCCGGTAACTCTGTTAGTCCAAACCGGTAAAGACACACTTATCCCACT 2259
2216 Db CGCCTTATCCGGTAACTCTGTTAGTCCAAACCGGTAAAGACACACTTATCCCACT 2275
2260 QY GGCAGCAGCCACTGGTAAACAGGATTAGCAGAGCAGGTATGTAGCGGTGCTTACAGAGTT 2319
2276 Db GGCAGCAGCCACTGGTAAACAGGATTAGCAGAGCAGGTATGTAGCGGTGCTTACAGAGTT 2335
2320 QY CTTGAAGTGGTGGCGCTTAACGCGCTACACTAGAGGACAGTATTTGGTATCTCGGCTCT 2379

Db 2336 CTTGAAAGTGGTGGCCCTAACTACCGCTACACTAGAGGACAGTATTTGGTATCTCGCGCTCT 2395
QY 2380 GCTGAAGCCAGTTACCTTTCGGAAGAGAGTGGTAGCTCTTGATCCGGCAAAACAAACCCAC 2439
Db 2396 GCTGAAGCCAGTTACCTTTCGGAAGAGAGTGGTAGCTCTTGATCCGGCAAAACAAACCCAC 2455
QY 2440 CGCTGTAGCGGTGGTTTTTTTTCGCAAGCAGCAGATTAACCGCAGAGAAAAAAGGATC 2499
Db 2456 CGCTGTAGCGGTGGTTTTTTTTCGCAAGCAGCAGATTAACCGCAGAGAAAAAAGGATC 2515
QY 2500 TCAAGAGAGATCCTTTTCATCTTTTCTACGGGTCTGACGCTCAGTGAACGAAACCTCAGC 2559
Db 2516 TCAAGAGAGATCCTTTTCATCTTTTCTACGGGTCTGACGCTCAGTGAACGAAACCTCAGC 2575
QY 2560 TTAAGGGATTTTGGTCATCAGATTAATCAAAAGGATCTTCACTAGATCTTTTAAATTA 2619
Db 2576 TTAAGGGATTTTGGTCATCAGATTAATCAAAAGGATCTTCACTAGATCTTTTAAATTA 2635
QY 2620 AAAATGAAGTTTAAATCAATCTAAAGTATATATAGTAAACCTTGTCTGACAGATTACCA 2679
Db 2636 AAAATGAAGTTTAAATCAATCTAAAGTATATATAGTAAACCTTGTCTGACAGATTACCA 2695
QY 2680 ATGCTTAATCAGTGAAGCCATCTCAGCGATCTCTATTTTCGTTTCATCCATAGTTGC 2739
Db 2696 ATGCTTAATCAGTGAAGCCATCTCAGCGATCTCTATTTTCGTTTCATCCATAGTTGC 2755
QY 2740 CTGACTCCCGCTCGTGTAGATACTACGATACGGAGGGCTTACCATCTGGCCCCAGTGC 2799
Db 2756 CTGACTCCCGCTCGTGTAGATACTACGATACGGAGGGCTTACCATCTGGCCCCAGTGC 2815
QY 2800 TGAATGATATCCGCGAGACCCAGCTCAGCGGCTCCAGATTTATCAGCAATAAACAGCC 2859
Db 2816 TGAATGATATCCGCGAGACCCAGCTCAGCGGCTCCAGATTTATCAGCAATAAACAGCC 2875
QY 2860 AGCCGAAGGGCCGACGCGAGAGTGGTCTGCAACTTTTATCCGCTCCATCCAGTCTAT 2919
Db 2876 AGCCGAAGGGCCGACGCGAGAGTGGTCTGCAACTTTTATCCGCTCCATCCAGTCTAT 2935
QY 2920 TAAATCTCCGGAAGCTAGAGTAAAGTATGTCGCCAGTTAATAGTTTGCAGAACCTGT 2979
Db 2936 TAAATCTCCGGAAGCTAGAGTAAAGTATGTCGCCAGTTAATAGTTTGCAGAACCTGT 2995
QY 2980 TGCATTTGTACAGGATCGTGGTGTACGCTCGTCTGGTATGCGTTTCAATTCAGCTC 3039
Db 2996 TGCATTTGTACAGGATCGTGGTGTACGCTCGTCTGGTATGCGTTTCAATTCAGCTC 3055
QY 3040 CGGTTCCCAACGATCAAGCGAGTTACATGATCCCCCATGTTGTGCAAAAAGCGGTAG 3099
Db 3056 CGGTTCCCAACGATCAAGCGAGTTACATGATCCCCCATGTTGTGCAAAAAGCGGTAG 3115
QY 3100 CTCCTTTCGCTCCGATCGTGTCAAGAGTAAAGTTGGCGAGTGTATCACTCATGT 3159
Db 3116 CTCCTTTCGCTCCGATCGTGTCAAGAGTAAAGTTGGCGAGTGTATCACTCATGT 3175
QY 3160 TATGGCAGCATGTCATAATCTCTTACTGTATGCCATCCGTAAGATGCTTTTCTGTGAC 3219
Db 3176 TATGGCAGCATGTCATAATCTCTTACTGTATGCCATCCGTAAGATGCTTTTCTGTGAC 3235
QY 3220 TGTGTAGTACTCAACCAAGTCAATCTGTAGATAGTGTATGGCGGACCGAGTGTCTCTTG 3279
Db 3236 TGTGTAGTACTCAACCAAGTCAATCTGTAGATAGTGTATGGCGGACCGAGTGTCTCTTG 3295
QY 3280 CCGCGGTCAATACGGGATAATACCGCCACATAGCAGAACTTTTAAAGTGTCTCATCAT 3339
Db 3296 CCGCGGTCAATACGGGATAATACCGCCACATAGCAGAACTTTTAAAGTGTCTCATCAT 3355
QY 3340 TGAAGAACGTTCTTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTCAGATCCAGTTC 3399
Db 3356 TGAAGAACGTTCTTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTCAGATCCAGTTC 3415
QY 3400 GATGTAAACCATCTGTGTGACCCCACTGATCTTTCAGCATCTTTTACCTTTTCAACGAGCTTC 3459

Db 3416 GATGTAAACCATCTGTGTGACCCCACTGATCTTTCAGCATCTTTTACCTTTCAACGCGTTTC 3475
QY 3460 TGGGTGAGCAAAAAACAGGAAGCAAAATGCCGCAAAAAAGGAATTAAGGGCGACACGGAA 3519
Db 3476 TGGGTGAGCAAAAAACAGGAAGCAAAATGCCGCAAAAAAGGAATTAAGGGCGACACGGAA 3535
QY 3520 ATGTTGAATACATCATCTCTTCTTTTCAATATTAATGAAGCATTTATCAGGGTTATTG 3579
Db 3536 ATGTTGAATACATCATCTCTTCTTTTCAATATTAATGAAGCATTTATCAGGGTTATTG 3595
QY 3580 TCTCATGACGGATACATATTTGAATGATTTAGAAAAATAAAACAATAGGGGTTCCGCG 3639
Db 3596 TCTCATGACGGATACATATTTGAATGATTTAGAAAAATAAAACAATAGGGGTTCCGCG 3655
QY 3640 CACATTTCCCGAAAGTGCAC 3662
Db 3656 CACATTTCCCGAAAGTGCAC 3678

RESULT 7
A20700
LOCUS A20700 5314 bp DNA linear PAT 22-AUG-1994
DEFINITION DMV1 DNA sequence.
ACCESSION A20700
VERSION A20700.1 GI:583402
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 5314)
AUTHORS
TITLE CARRIER-BOUND RECOMBINANT PROTEIN, PROCESS FOR PRODUCING IT AND ITS
JOURNAL USE AS AN IMMUNOGEN AND VACCINE
Patent: WO 9113155-A 4 05-SEP-1991;
FEATURES
source
1..5314
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"

ORIGIN

Query Match 71.6%; Score 2623.6; DB 6; Length 5314;
Best Local Similarity 83.4%; Pred. No. 0;
Matches 3059; Conservative 0; Mismatches 579; Indels 28; Gaps 6;

QY 20 ATTTTGTAAAATTCGCGTTAAATTTTGTAAATCAGCTCAATTTTAAACCAATAGGCC 79
Db 1651 ATTTTGTAAAATTCGCGTTAAATTTTGTAAATCAGCTCAATTTTAAACCAATAGGCC 1710
QY 80 GAAATCGGCAAAATCCCTTATAAATCAAAAGAAATAGACCGAGATAGGGTTGAGTGTGT 139
Db 1711 GAAATCGGCAAAATCCCTTATAAATCAAAAGAAATAGACCGAGATAGGGTTGAGTGTGT 1770
QY 140 CCAGTTTGAACAAGATCCACTATTAAGACGCTGACCTCCAACTCAAGGGCGAAAA 199
Db 1771 CCAGTTTGAACAAGATCCACTATTAAGACGCTGACCTCCAACTCAAGGGCGAAAA 1830
QY 200 ACCGCTATACGCGGATGCGCCACTACGCTGAACCATCACTCAAGTTTTCGGGG 259
Db 1831 ACCGCTATACGCGGATGCGCCACTACGCTGAACCATCACTCAAGTTTTCGGGG 1890
QY 260 TCGAGGTGCGTAAAGCATAAATCGGAACCCCTAAAGGGAGGCCCCCGATTTTAGAGCTTGA 319
Db 1891 TCGAGGTGCGTAAAGCATAAATCGGAACCCCTAAAGGGAGGCCCCCGATTTTAGAGCTTGA 1950
QY 320 CCGGGAAAGCCCGCGAAGTGGCGAGAAAGGAAGGAAAGGAAAGGAGCGGCGGCT 379
Db 1951 CCGGGAAAGCCCGCGAAGTGGCGAGAAAGGAAGGAAAGGAGCGGCGGCT 2010
QY 380 AGGGCGCTGGCAAGTGTAGCGGTACGCTGCGGTAAACCAACACCCCGCGCTTAAT 439
Db 2011 AGGGCGCTGGCAAGTGTAGCGGTACGCTGCGGTAAACCAACACCCCGCGCTTAAT 2070

QY 440 GGGCGCTACAGGCGCGTCCCAATTCGCAATTCAGGCTCGCAATCTGTTGGGAAGGGCA 499
Db 2071 GCGCGCTACAGGCGCGTCCCAATTCGCAATTCAGGCTCGCAATCTGTTGGGAAGGGCA 2130
QY 500 TCGGTGCGGCGCTCTTCGCTATTAACGCAAGCTGGCGAAGGGGATGCTCTCAAGGGCA 559
Db 2131 TCGGTGCGGCGCTCTTCGCTATTAACGCAAGCTGGCGAAGGGGATGCTCTCAAGGGCA 2190
QY 560 TTAAGTTGGTAAAGCGCACGGTTTCCAGTCACGAGCTTGGAACGAGCGCCAGTGAG 619
Db 2191 TTAAGTTGGTAAAGCGCACGGTTTCCAGTCACGAGCTTGGAACGAGCGCCAGTGAG 2250
QY 620 CGCGCTCGTTCATTCACGCTTTTGAACCCGTTGGAGACCGGCAGACTCGCGTGCACAT 679
Db 2251 TTGTAATACGACTCACTATAGGCGGAATTGGAGCTCCACCGCGTGGCGCGCTCTAGT 2310
QY 680 GTGTTTACAGCGTGAATGGAGCAGATGAAGATGCTCGACACGCTGGACAACAGCGAGCTA 739
Db 2311 ATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGATAGTTAAGCCAGTATATACACTC 2370
QY 740 GATTAAACCTAGAAAGATAATCATATTGTGACGTAGTTAAAGATAATCATCGTAAAT 799
Db 2371 CGCTATCGTACGACTGGGCTGCGGCTGCGCCCGACACCCGCCAACCACCGCTGAGC 2430
QY 800 TGACGCTGGGATCTGTAATACGACTCACTATAGGGCGAATTTGGGTACCGGGCCCCCCT 859
Db 2431 CGCCCTGACGGCTTGTCTGCTCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCG 2490
QY 860 CGAGTTCGACGATTCGA-----TAAGCTTGATTCGAATTCGAGCTCGACCCCGGGGATC 913
Db 2491 GGAGCTGCAATGCTCAGAGGTTTACCCGTCATCACCGAAGCGCGAGGAGTAAAGTTC 2550
QY 914 CACTAGTTCTAGAGCGGCCACCGCGTGGAGCTCCAGCTTTTCTCCCTTTAGTGAG 973
Db 2551 GGATGCTTTGTGAGCAATTCGCTTAAAGTAAGCAATTCGTTAAAGTCTGACTGTGC 2610
QY 974 GGTAAATAGATCCCAT-----GCGTCAATTTTACGAGACTATCTTTCTAGGGTTAA 1026
Db 2611 GGATCACCGCTTCCAGTAGCGACAGAACAAATTGATTTGTTAAATTTTCGAGAGAAAGATCG 2670
QY 1027 TCTAGTGCATCAGGATCATATCTGTCGGGTCTTTTTCGGCTCAGTCATCGCCCAAGCT 1086
Db 2671 CGAGGAAGATCAATACATAAAGAGTTGAATCTTTTGTGTTCTCGACATGGGTAAATCCT 2730
QY 1087 GCGCTATCTGCGCATCGGGAGGAAGACCGTGCCTTTTCCGCGAGGTTGAAGCG 1146
Db 2731 CATGTTTGAATGCGCTAGAGATCGGCAAGCTTGCAATGCTGAGCTGCACTCTAGA 2790
QY 1147 CATGAAAGAGTTTCCGAGGATGACTGCTGTCATTGACGTTGAGCGTGAAGAAACGACGT 1206
Db 2791 GGATCCCGACGCTCGAGCCATTAATATGTTTTCGGTAAATTCAGCGCTTCCATGAT 2850
QY 1207 TTACCATGATGATTCGGGAAGTGTGGCCATCGACGCTTTAAACGGTGAACCTGTTGCTTC 1266
Db 2851 GAGACAGGCGGTTT--GAATGTTGAGCGGATGAACATATAAGCAATGACGGCAGCAATA 2908
QY 1267 AGGCCACCTGGGATACCACTGCTCGCGCTTTTCCGACACAGTTCCGGATGCTCAGCG 1326
Db 2909 AACTCAACAGG---AGCAAGAAAGCGGGATTCACCAAAATCGCAGCTACCAATAAGC 2965
QY 1327 CGAAGCGCATCAGCAACCGCAATACCGGCGACAG-----CGGAACTGCCGTGCCG 1380
Db 2966 CAAGCTCAACCGACGACGACGAGAGCGGTGAGTAGCAATCCAACTTTGTTACTC 3025
QY 1381 GTGTGAGATTAATGACAGCGGTGCGGCTGGGATATTAAGTTCAGCGAGGACGGGTAT- 1439
Db 3026 GTCAGAAATCGAAATCATCTTCGTTTAAATCCAAACCGGCAAGACCTGAAATGAAAT 3085
QY 1440 --CTTGGCTGATGCGCGAGAAATGGAATGATACCCCGTGAAGTTACCGCGCGCGC 1496
Db 3086 CGACCTCGAGGGGGCGGATCCGACGCTTTTGTCCCTTTAGTGAGGTTAATTCGA 3145
QY 1497 GCTTGGCGTAATCATGCTCATAGCTGTTTCTGTGTGAATTTGTTATCCGCTCACAATTC 1556

Db 3146 GCTTGGCGTAATCATGCTCATAGCTGTTTCTGTGTGAATTTGTTATCCGCTCACAATTC 3205
QY 1557 CACACAAATACAGCGCGGAGCATAAAGAGCTTAAAGCCTGGGGTCCCTAATCAGTGAGCT 1616
Db 3206 CACACAAATACAGCGCGGAGCATAAAGAGCTTAAAGCCTGGGGTCCCTAATCAGTGAGCT 3265
QY 1617 AACTCACATTAATTCGCTTGCCTCACTCGCCGCTTTTCCAGTCGGGAAACCTGTGCTGCC 1676
Db 3266 AACTCACATTAATTCGCTTGCCTCACTCGCCGCTTTTCCAGTCGGGAAACCTGTGCTGCC 3325
QY 1677 AGCTGCAATTAATGAATTCGCGCAACCGCGGGGAGAGCGGTTTGGGTATTTGGCGCTCTT 1736
Db 3326 AGCTGCAATTAATGAATTCGCGCAACCGCGGGGAGAGCGGTTTGGGTATTTGGCGCTCTT 3385
QY 1737 CGGCTTCCCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCG 1796
Db 3386 CGGCTTCCCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCG 3445
QY 1797 CTCACCTCAAAGCGCTTAATACGCTTATCCACAGAAATCAGGGGATAAACGAGGAAAGACA 1856
Db 3446 CTCACCTCAAAGCGCTTAATACGCTTATCCACAGAAATCAGGGGATAAACGAGGAAAGACA 3505
QY 1857 TGTGAGCAAAAGGCGCAGCAAAAGGCGAGAACCGGTAAAGGCGCGGTTGCTGGCGTTTT 1916
Db 3506 TGTGAGCAAAAGGCGCAGCAAAAGGCGAGAACCGGTAAAGGCGCGGTTGCTGGCGTTTT 3565
QY 1917 TCCATAGGCTCGCCCGCTGACGAGCATCAAAAATTCGACGCTCAAGTCAGAGGTGCG 1976
Db 3566 TCCATAGGCTCGCCCGCTGACGAGCATCAAAAATTCGACGCTCAAGTCAGAGGTGCG 3625
QY 1977 GAAACCCGACAGGACTATAAAGATACAGGCGCTTTCCCTCGAAGCTCCCTCGTGCCT 2036
Db 3626 GAAACCCGACAGGACTATAAAGATACAGGCGTTCCCTCGAAGCTCCCTCGTGCCT 3685
QY 2037 CTCCTGTTTCGACCTCGCGCTTACCGGATACGCTCGCGCTTTTCCCTTCGGGAAGCG 2096
Db 3686 CTCCTGTTTCGACCTCGCGCTTACCGGATACGCTCGCGCTTTTCCCTTCGGGAAGCG 3745
QY 2097 TGGCGCTTTCATAGCTCAGCTGCTAGTATCTCAGTTTCGGTGTAGTTCGTTCCGTTCA 2156
Db 3746 TGGCGCTTTCATAGCTCAGCTGCTAGTATCTCAGTTTCGGTGTAGTTCGTTCCGTTCA 3805
QY 2157 AGCTGGGCTGTGTGACGAAACCCCGCTTCAGCGGACCGCTGCGCTTATCCGCTAACT 2216
Db 3806 AGCTGGGCTGTGTGACGAAACCCCGCTTCAGCGGACCGCTGCGCTTATCCGCTAACT 3865
QY 2217 ATCGCTTTCGAGTCAACCGGTAAGACAGACTTATCGGCACTGGCAGCAGCACTGCTA 2276
Db 3866 ATCGCTTTCGAGTCAACCGGTAAGACAGACTTATCGGCACTGGCAGCAGCACTGCTA 3925
QY 2277 ACAGGATTAGCAGAGGAGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTTA 2336
Db 3926 ACAGGATTAGCAGAGGAGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTTA 3985
QY 2337 ACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCAGTTACCT 2396
Db 3986 ACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCAGTTACCT 4045
QY 2397 TCGGAAAAAGAGTGGTAGCTCTTGAATCCGGCAAAACAAACCGCTGGTGGTGGTGGT 2456
Db 4046 TCGGAAAAAGAGTGGTAGCTCTTGAATCCGGCAAAACAAACCGCTGGTGGTGGTGGT 4105
QY 2457 TTTTGTGTTGCAAGCAGCAGATTAACCGCAGAAAAAAGGATCTCAAGAGATCTTTTCA 2516
Db 4106 TTTTGTGTTGCAAGCAGCAGATTAACCGCAGAAAAAAGGATCTCAAGAGATCTTTTCA 4165
QY 2517 TCTTTTCTACGGGCTGACGCTCAGTGAGAACGAAAACTCACTTGAAGGATTTTGGTCA 2576
Db 4166 TCTTTTCTACGGGCTGACGCTCAGTGAGAACGAAAACTCACTTGAAGGATTTTGGTCA 4225
QY 2577 TGAGATTAATCAAAAAGGATCTTCACTAGATCTTTTAAATTTAAAAATGAAGTTTAAAT 2636

Dd	2311	ATGGTGCACTCTCAGTACAACTGCTCTGATGCCGATAGTTAAGCCAGTATATACACTC	2370
Qy	740	GATTAAACCTCAGAAAGATAATCATATTTGACGTACGTTTAAAGATAATCATCGTAAAT	799
Dd	2371	CGCTATCGGTACGTGACTGGGTCAATGGCTGCGCCCGACACCCGCGCAACCCCGCTGACG	2430
Qy	800	TGACGCATGGGATCTGTATACGACTCACTATAGGCGGAATTTGGGTACCGGGCCCCCCT	859
Dd	2431	CGCCCTGACGGCTGTCTGCTCCGGCATCCGCTTACAGACAAGCTGTGACGCTCTCCG	2490
Qy	860	CGAGGTGACGGTATCGA-----TAAGCTTGATATCGAATTCCTGCAGCCCGGGGATC	913
Dd	2491	GGAGCTGATGTGTGACAGGTTTTCACCGTCAATCACCGAAACCGCGAGGCAGTAAAGTTC	2550
Qy	914	CACCTAGTTCTAGAGGGCGCCACCGCGGTGAGCTCCAGCTTTTGTCCCTTTAGTGAG	973
Dd	2551	GGATGCTTTGTGAGCAATTCGTCCTTAAAGTAAGCAATTCGCTGTAAGTGTCACTGTGC	2610
Qy	974	GGTTAATTAGATCCCAT-----GCGTCAATTTTACGCAATATCTTTCTAGGGTTAA	1026
Dd	2611	GGATCACCGCTCCAGTAGCGACAGAAGCAATTGATTGGTAAATTTGAGAGAAAGATCG	2670
Qy	1027	TCTAGTGCATCAGGATCATATCTGCGGTCTTTTTCGGCTCAGTCAATCCGCCAAGCT	1086
Dd	2671	CGAGGAAGATCAATACATAAAGAGTTGAATCTTTTGTGTCTTCGACATGGGTAACTCT	2730
Qy	1087	GCGCTATCTGGGCATCGGGAGGAAGAGCCGCTGCTTTTCCGCGAGGTTGAAGCGG	1146
Dd	2731	CATGTTTGAATGGCCCTAGAGATCGGCAAGCTTGATGCTGCAAGTCACTCTAGA	2790
Qy	1147	CATGGAAGAGTTTGC CGAGGATGACTGCTGCTGCATTGACGTTGAGCGAATAACGACGT	1206
Dd	2791	GGATCCCGACGCTCGACGCCAATAATATGTTTTCGGTAAATTCAGCGCCTTCATGAT	2850
Qy	1207	TTACATGATGATTCGGAGAGTGTGGCCATGCAAGCCCTTTAAACGGTGAATGTTGCTTC	1266
Dd	2851	GAGACAGGCGGTTT--GAATGTTGAGCGGATGAACATATAAGCAATGACGCGACCAATA	2908
Qy	1267	AGGCCACCTGGGATACCACTGCTGCGGCTTTTCCGGACACAGTTTCCGGATGGTCAGCC	1326
Dd	2909	AACTCAACAGG---AGCAGGAAGCGAGGATCCCAAAAGTCCAGCGTACCATAAAGC	2965
Qy	1327	CGAAGCGCATCAGCAACCCGGAACAATACCGGGACAG-----CCGGAACCTCCCGTCCG	1380
Dd	2966	CAAGCTCAACGACGACGACGACGAGAGCGGTCACTAGTAAATCCAAACTTTGTACTC	3025
Qy	1381	GTGTGAGATTAATGACAGCGGTGCGCGCTGGGATATTACGTGACGAGGACGGGTAT-	1439
Dd	3026	GTCAAGAAATCGAAATCATCTTTCGGTTAAATCCAAACCGGCAAGACCTGAATGAGAAAT	3085
Qy	1440	---CCTGGCTGATGCCGAGAAATGGACATGGATACCCCGTGGATTACC CGCGGGCGC	1496
Dd	3086	CGACCTCGAGGGGGGCCGCTACCGAGCTTTTGTCCCTTTTAGTGAGGTTAATTCGA	3145
Qy	1497	GCTTGGCGTAAATCATGGCTAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCACAATTC	1556
Dd	3146	GCTTGGCGTAAATCATGGTCAATAGCTTTCTCTGTGTGAAATTTGTTATCCGCTCACAATTC	3205
Qy	1557	CACACAACATACGACCGGAGCATAAAGTGTAAAGCTTGGGCTGCGCTTAATCAGTGAGCT	1616
Dd	3206	CACACAACATAGGAGCGGAAGCATAAAGTGTAAAGCTTGGGCTGCGCTTAATCAGTGAGCT	3265
Qy	1617	AACTCACATTAATTTGGTTGGCTCACTGCCGCTTTCCAGTCGGGAACCTGTGCTGCC	1676
Dd	3266	AACTCACATTAATTTGGTTGGCTCACTGCCGCTTTCCAGTCGGGAACCTGTGCTGCC	3325
Qy	1677	AGCTGATTAATTAATTCGGCCAAACGCGCGGGGAGAGCGGCTTTTGGGTATTGGCGCTCTT	1736
Dd	3326	AGCTGATTAATTAATTCGGCCAAACGCGCGGGGAGAGCGGCTTTTGGGTATTGGCGCTCTT	3385
Qy	1737	CGGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGGTTCGGCTGCGGCGAGCGGTATCAG	1796
Dd	3386	CGGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGGTTCGGCTGCGGCGAGCGGTATCAG	3445
Qy	1797	CTCACTCAAAGCGGTAAATACGGTTATCCACAGAATCAGGGGATAAACGAGGAAAGAAACA	1856
Dd	3446	CTCACTCAAAGCGGTAAATACGGTTATCCACAGAATCAGGGGATAAACGAGGAAAGAAACA	3505
Qy	1857	TGTGAGCAAAAAGCGCAGCAAAAAGCGCAGAAACCGGTAAAAAAGCGCGGTTCGTGGCGTTT	1916
Dd	3506	TGTGAGCAAAAAGCGCAGCAAAAAGCGCAGAAACCGGTAAAAAAGCGCGGTTCGTGGCGTTT	3565
Qy	1917	TCCATAGGCTCGCCCCCTGACGAGCATCAAAAATTCGACGCTCAAGTCAGAGGTGCG	1976
Dd	3566	TCCATAGGCTCGCCCCCTGACGAGCATCAAAAATTCGACGCTCAAGTCAGAGGTGCG	3625
Qy	1977	GAAAACCCGACAGGACTATAAAGATACCAAGCGCTTTCCCCCTGGAAGCTCCCTCGTGCCT	2036
Dd	3626	GAAAACCCGACAGGACTATAAAGATACCAAGCGCTTTCCCCCTGGAAGCTCCCTCGTGCCT	3685
Qy	2037	CTCCTGTTCCGACCCCTGCGCTTACCGGATACCTGTCCGCTTCTCCCTTCGGGAAGCG	2096
Dd	3686	CTCCTGTTCCGACCCCTGCGCTTACCGGATACCTGTCCGCTTCTCCCTTCGGGAAGCG	3745
Qy	2097	TGGCGCTTCTCATAGCTCAGCTGTAGGTATCTCAGTTTCGGGTGAGGTCTTCGCTCCA	2156
Dd	3746	TGGCGCTTCTCATAGCTCAGCTGTAGGTATCTCAGTTTCGGGTGAGGTCTTCGCTCCA	3805
Qy	2157	AGCTGGGCTGTGTGACACGAAACCCCGCTTTCAGCCCGACCGCTGCGCTTATCCGGTAACT	2216
Dd	3806	AGCTGGGCTGTGTGACACGAAACCCCGCTTTCAGCCCGACCGCTGCGCTTATCCGGTAACT	3865
Qy	2217	ATCGCTTTGAGTCCAAACCCGTTAAGACAGCTTATCGGCACCTGGCAGCAGCACCTGGTA	2276
Dd	3866	ATCGCTTTGAGTCCAAACCCGTTAAGACAGCTTATCGGCACCTGGCAGCAGCACCTGGTA	3925
Qy	2277	ACAGATTTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGTGSCCTA	2336
Dd	3926	ACAGATTTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGTGSCCTA	3985
Qy	2337	ACTACGGCTTACTAGAAAGGACAGTATTTGGTATCTCGGCTCTGCTGTAAGCAGTTACCT	2396
Dd	3986	ACTACGGCTTACTAGAAAGGACAGTATTTGGTATCTCGGCTCTGCTGTAAGCAGTTACCT	4045
Qy	2397	TCGGAAGAGAGTGTGTAGCTTTCATCCGCAAAACAAACACCCGCTGCTAGCGGTGTT	2456
Dd	4046	TCGGAAGAGAGTGTGTAGCTTTCATCCGCAAAACAAACACCCGCTGCTAGCGGTGTT	4105
Qy	2457	TTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAGATCTCTTCA	2516
Dd	4106	TTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAGATCTCTTCA	4165
Qy	2517	TCTTTTCTACGGGCTCTGACGCTCAGTGGAAACGAAAACTCAGTTAAGGGATTTTGGTCA	2576
Dd	4166	TCTTTTCTACGGGCTCTGACGCTCAGTGGAAACGAAAACTCAGTTAAGGGATTTTGGTCA	4225
Qy	2577	TGAGATTAACAAAAAGGATCTTTCACCTAGTCTTTTAAATTTAAAAATGAAGTTTAAAT	2636
Dd	4226	TGAGATTAACAAAAAGGATCTTTCACCTAGTCTTTTAAATTTAAAAATGAAGTTTAAAT	4285
Qy	2637	CAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTTACCAATGCTTAAATCAGTGAGG	2696
Dd	4286	CAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTTACCAATGCTTAAATCAGTGAGG	4345
Qy	2697	CACCTATCTCAGCGATCTGTCTATTTTGGTTCATTCATAGTTGCTGACTCCCGCTGCT	2756
Dd	4346	CACCTATCTCAGCGATCTGTCTATTTTGGTTCATTCATAGTTGCTGACTCCCGCTGCT	4405
Qy	2757	AGATAACTACGATACGGGAGGCTTACCATCTGCGGCCAGTCTGCTCAATGATACCGCGAG	2816
Dd	4406	AGATAACTACGATACGGGAGGCTTACCATCTGCGGCCAGTCTGCTCAATGATACCGCGAG	4465
Qy	2817	ACCCACGCTCACCGGCTCCAGATTTTATCAGCAATAAAACCCAGCAGCCGGAAGGCGCGAGC	2876
Dd	4466	ACCCACGCTCACCGGCTCCAGATTTTATCAGCAATAAAACCCAGCAGCCGGAAGGCGCGAGC	4525

```
QY 2877 GCAGAGTGGTCTGCAACTTATCCGCTCCATCCAGTCTATTAAATGTTGCGCGAAG 2936
Db 4526 GCAGAGTGGTCTGCAACTTATCCGCTCCATCCAGTCTATTAAATGTTGCGCGAAG 4585
QY 2937 CTAGAGTAAAGTAGTTCGCGCACTTAATAGTTTGGCAACGTTGTTGCCATTGCTACAGGCA 2996
Db 4586 CTAGAGTAAAGTAGTTCGCGCACTTAATAGTTTGGCAACGTTGTTGCCATTGCTACAGGCA 4645
QY 2997 TCGTGGTGCAGCTCGTCTGTTGGTATGCGTTTCATTTCAGTCCGTTCCCAACGATCAA 3056
Db 4646 TCGTGGTGCAGCTCGTCTGTTGGTATGCGTTTCATTTCAGTCCGTTCCCAACGATCAA 4705
QY 3057 GCGAGTTACATCATGATCCCATGTTGTCGCAAAAAGGTTAGTCTTGGTCTCCGCA 3116
Db 4706 GCGAGTTACATCATGATCCCATGTTGTCGCAAAAAGGTTAGTCTTGGTCTCCGCA 4765
QY 3117 TCGTTCAGAAAGTAAAGTTGGCGCAGTGTATCACTCATGTTATGGCAGCACTCATAT 3176
Db 4766 TCGTTCAGAAAGTAAAGTTGGCGCAGTGTATCACTCATGTTATGGCAGCACTCATAT 4825
QY 3177 ATTCTCTTACTGTCAATGCCATCGTAAAGATGCTTTTCTGTGACTGGTGAAGTACTCAACCA 3236
Db 4826 ATTCTCTTACTGTCAATGCCATCGTAAAGATGCTTTTCTGTGACTGGTGAAGTACTCAACCA 4885
QY 3237 AGTCATTCTGAGAAATAGTATGCGCGCAGGAGTTGCTCTTCCCGCGGTCAATACGGG 3296
Db 4886 AGTCATTCTGAGAAATAGTATGCGCGCAGGAGTTGCTCTTCCCGCGGTCAATACGGG 4945
QY 3297 ATAAATACCGCGCACATAGCAGAACTTTAAAGTGTCTCATCTGCAAAAAGGTTCTTCGG 3356
Db 4946 ATAAATACCGCGCACATAGCAGAACTTTAAAGTGTCTCATCTGCAAAAAGGTTCTTCGG 5005
QY 3357 GCGCAAACTCTCAAGGATCTTACCGCTGTGAGATCCAGTTCGATGTAAACCCACTCGTG 3416
Db 5006 GCGCAAACTCTCAAGGATCTTACCGCTGTGAGATCCAGTTCGATGTAAACCCACTCGTG 5065
QY 3417 CACCCAACTGATCTTCAAGATCTTTTACCTTACAGAGGTTTCTGGTGAGCAAAAACAG 3476
Db 5066 CACCCAACTGATCTTCAAGATCTTTTACCTTACAGAGGTTTCTGGTGAGCAAAAACAG 5125
QY 3477 GAAGGCAAAATGCGCAAAAAGGGAATAAGGCGCAGACGGAATGTTGAATCTCATAC 3536
Db 5126 GAAGGCAAAATGCGCAAAAAGGGAATAAGGCGCAGACGGAATGTTGAATCTCATAC 5185
QY 3537 TCTTCCTTTTCAATATTTAAGCAATTTATCAGGGTTATGTCATCAGCGGATACA 3596
Db 5186 TCTTCCTTTTCAATATTTAAGCAATTTATCAGGGTTATGTCATCAGCGGATACA 5245
QY 3597 TATTTGAATGTTATGAAAAATAAACAATAGGGGTTCCGGCGACATTTCCCGAAAAAG 3656
Db 5246 TATTTGAATGTTATGAAAAATAAACAATAGGGGTTCCGGCGACATTTCCCGAAAAAG 5305
QY 3657 TGCCAC 3662
Db 5306 TGCCAC 5311
```

```
RESULT 9
AF171227/c
LOCUS AF171227 3552 bp DNA circular SYN 04-FEB-2000
DEFINITION Shuttle vector pSLIRES11, complete sequence.
ACCESSION AF171227
VERSION AF171227.1 GI:5764544
KEYWORDS
SOURCE
ORGANISM Shuttle vector pSLIRES11
other sequences; artificial sequences; vectors.
REFERENCE 1 (bases 1 to 3552)
AUTHORS Chen, C.M., Smith, D.M., Peters, M.A., Samson, M.E., Zitz, J.,
Tabin, C.J., and Cepko, C.L.
TITLE Production and design of more effective avian
replication-incompetent retroviral vectors
JOURNAL Dev. Biol. 214 (2), 370-384 (1999)
```

```
99456779 MEDLINE
10525341 PUBMED
2 (bases 1 to 3552) REFERENCE
Chen, C.-M.A., Samson, M.E.S. and Cepko, C.L. AUTHORS
Direct Submision TITLE
Submitted (23-JUL-1999) Genetetics, Harvard Medical School, 200 JOURNAL
Longwood Avenue, Boston, MA 02115, USA
FEATURES
source
1..3552
/organism="Shuttle vector pSLIRES11"
/mol_type="genomic DNA"
/db_xref="taxon:101936"
/lab_host="Escherichia coli"
join(2864..3552,1..2208)
note="similar to pBluescript"
2236..2789
note="IRES of EMCV"
2790..2849
note="multiple cloning site"
ORIGIN
Query Match 71.3%; Score 2610.8; DB 12; Length 3552;
Best Local Similarity 83.8%; Pred. No. 0;
Matches 3070; Conservative 0; Mismatches 482; Indels 110; Gaps 6;
QY 1 CTAATTTGTAAGCGTTAAATTTTGTAAATTCGCGTTAAATTTTGTAAATTCAGCTC 60
Db 3552 CTAATTTGTAAGCGTTAAATTTTGTAAATTCGCGTTAAATTTTGTAAATTCAGCTC 3493
QY 61 ATTTTAAACCAATAGGCGCAAAATCGGCAAAATCCCTTAAATCAAAAGAAATAGACCGA 120
Db 3492 ATTTTAAACCAATAGGCGCAAAATCGGCAAAATCCCTTAAATCAAAAGAAATAGACCGA 3433
QY 121 GATAGGTTGAGTGTGTTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGAAACGTGGACTC 180
Db 3432 GATAGGTTGAGTGTGTTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGAAACGTGGACTC 3373
QY 181 CAACGTCAAAAGGCGCAAAAACCGTCTATCAGGCGATGCGCCCACTACGTGAACCATCACC 240
Db 3372 CAACGTCAAAAGGCGCAAAAACCGTCTATCAGGCGATGCGCCCACTACGTGAACCATCACC 3313
QY 241 CTAATCAAGTTTGGGTCGAGTCCGTAAGCACTTAAATCGGAACCTTAAGGGAG 300
Db 3312 CTAATCAAGTTTGGGTCGAGTCCGTAAGCACTTAAATCGGAACCTTAAGGGAG 3253
QY 301 CCCCCGATTTAGAGCTTGACGCGGAAAGCCGCGAAGCTGGCGAGAAAGGAAGGAAGAA 360
Db 3252 CCCCCGATTTAGAGCTTGACGCGGAAAGCCGCGAAGCTGGCGAGAAAGGAAGGAAGAA 3193
QY 361 AGCGAAAGAGCGGCGCTAGGCGCTGGCAAGTGTAGCGGTACCGCTGCGGTAAACCCAC 420
Db 3192 AGCGAAAGAGCGGCGCTAGGCGCTGGCAAGTGTAGCGGTACCGCTGCGGTAAACCCAC 3133
QY 421 CACACCCCGCGCTTAATGCGCCGCTACAGGCGGCTCCCATTCGCCATTTCAGGCTGG 480
Db 3132 CACACCCCGCGCTTAATGCGCCGCTACAGGCGGCTCCCATTCGCCATTTCAGGCTGG 3073
QY 481 CAACTGTTGGGAAGGCGGATCGGTGCGGCTCTTTCGCTATTACGCGAGCTGGCGAAAGG 540
Db 3072 CAACTGTTGGGAAGGCGGATCGGTGCGGCTCTTTCGCTATTACGCGAGCTGGCGAAAGG 3013
QY 541 GGGATGTGCTCAAGCGGATTAAGTTGGTAAACGCGAGGGTTTCCAGTTCAGGCTTG 600
Db 3012 GGGATGTGCTCAAGCGGATTAAGTTGGTAAACGCGAGGGTTTCCAGTTCAGGCTTG 2953
QY 601 TAAACGAGCGCGCAGTACGCGGCTCGCTTCATTTCAGTTTGTGAACCCGCGGAGACGG 660
Db 2952 TAAACGAGCGCGCAGTACGCGGCTCGCTTCATTTCAGTTTGTGAACCCGCGGAGACGG 2913
QY 661 GCAGACTCGCGTGCATAATGTGTTTACAGCGGTGATGGAGCAGATGAAGATGCTCGACAC 720
Db 2912 GCGGAATGGGTACCGGGCCCCCTCGATCGAGGTGAGGTATCGATAGCTTGGGCT 2853
```

QY 721 GCTGCAGAACCGCAGCTAGATAATACCCCTAGAAAGATAATCATATTTGTGACGTACGTTAA 780
|||
Db 2852 GCAGCG-----CCGCACTAGTGATTCGGATCCCGGCGAGCTCGAAT 2808
QY 781 AGATTAATCATGCGTAAATTTGACGATGGGATCTGTATATACGACTCATATAGGCGCGAAT 840
|||
Db 2807 TCGTAATCATGGCCATGATGTAATATCATCGTGTTTTTCAAAGGAAACCAACGCTCCCGGTGG 2748
QY 841 TGGGTACCGGGCCCCCTCGAGGTCGACGGTATCGATAAGCTTCGATATCGAATTCCTGC 900
|||
Db 2747 TTCGGGGGCCCTAGACGTTTTTAACTCGACTAAACACACGTAAGCATGTGCACCGAG 2688
QY 901 AGCCCGGGGGATPCACTAGTTCTTAGAGCGCGCCACCGCGGTGGAGCTCCAGCTTTTGT 960
|||
Db 2687 GCCCCAGATCAGATCCCATACAAATGGGTACCTTCTGGGCACTCTTCAGCCCTTTGTTGA 2628
QY 961 TCCCTTTAGTAGGGTTAAATAGATCCCATGGGTCAATTTTACGGCAGACTATCTTTCTAG 1020
|||
Db 2627 ATACGCTTTAGGAGAGCCATTTGACTTTTCACAACTATCCAACTCAACGTTGGCACT 2568
QY 1021 GGTAACTAGCTGATCAGGATCATATCGTGGGTCTTTTTCGGGCTCAGTCATCGCC 1080
|||
Db 2567 GG-----GGTTGCGCCCTTTGCGAGGTGATCTTATAACGTGGCTTTTGCC 2519
QY 1081 CAAGCTGGCGCTATCTGGGCATCGGGAGGAAGAACCGCGTGCCTTTTCCCGCAGGTTG 1140
|||
Db 2518 GCAGAGCACCTGTCCGACGGTGGGGTTCGCTGCTGCAAGGGTCTGTACAGACGT 2459
QY 1141 AAGCGCATGGAAGATTGCGGAGGATGACTGCTGCTGCATTTGACGTTGAGCGAAAC 1200
|||
Db 2458 TGTTTTGTCTCAAGAAGCTTCCAGAGGA-----ACTGCTCTTCCACGACATTCACAGA 2404
QY 1201 GCACGTTTACCATGATGATTCGGGAAGTGTGGCCATGCAGCCCTTTTAAAGGTGAAGTGT 1260
|||
Db 2403 CTTGCAATCTCTTTCGCGAGAGGGGAAGACCCCTAGGAATGCTCGTCAAGAGACAGG 2344
QY 1261 TGTGTTCAGGCCACCTGGGATACAGTTCTGTGCGGCTTTTCCGGACACAGTTTCCGGATGG 1320
|||
Db 2343 CCAGGTTTCGGGCTCTCACATTCGCAAAAGACGGCAATATGGTGGAAATAACATATAG 2284
QY 1321 TCAGCCGAGGCGCATCAGCAACCCGAAACATACCGGGACAGCCGGAACTGCCGTGGC 1380
|||
Db 2283 ACAACGCA-----CACCGGCTTATTCCAAGCGGCTTCGGCCAGTAAGCTTAGGGGG 2231
QY 1381 GTGTGCAATTAATGACAGCGGTGCGGCTGGGATATTACGTACGCGAGGACGGGTATC 1440
|||
Db 2230 GGGGAGGAGTCTAGAAATCGATGAGCTCCAGCTTTTGTTCCTTTAGTGGGTTAAT- 2172
QY 1441 CTGGCTGGATGCCGAGAAATGGAATGGAATACCCCGTGAAGTTACCCGCGGGCGCGCTT 1500
|||
Db 2171 -----TTCGAGCTT 2163
QY 1501 GCGTAATCATGGTCATAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCACAAATTCACAA 1560
|||
Db 2162 GCGTAATCATGGTCATAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCACAAATTCACAA 2103
QY 1561 CAACATAGAGCCGGAAGCATAAAGTGTAAAGCTGGGTGCTTAATGAGTGAGCTAACT 1620
|||
Db 2102 CAACATAGAGCCGGAAGCATAAAGTGTAAAGCTGGGTGCTTAATGAGTGAGCTAACT 2043
QY 1621 CACATTAATTTGGTTGCGCTCATCTGCCGCTTTCCAGTCGGGAAACCTGTGTCGACGT 1680
|||
Db 2042 CACATTAATTTGGTTGCGCTCATCTGCCGCTTTCCAGTCGGGAAACCTGTGTCGCACT 1983
QY 1681 GCATTAATGAATCGGCCAAACCGGCGGGGAGAGGGCGGTTTGCGTATTTGGCGCTCTTCGC 1740
|||
Db 1982 GCATTAATGAATCGGCCAAACCGGCGGGGAGAGGGCGGTTTGCGTATTTGGCGCTCTTCGC 1923
QY 1741 TTCCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGGCGAGCGGTATCAGCTCA 1800
|||
Db 1922 TTCCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGGCGAGCGGTATCAGCTCA 1863
QY 1801 CTCAAAGCGGTAATACCGGTTTATCCACAGAAATCAGGGGATTAACGCGAGGAAACATGTG 1860
|||

Db 1862 CTCAAAGCGGTAATACCGGTTATCCACAGAAATCAGGGATTAACGCGAGGAAAGACATGTG 1803
|||
QY 1861 AGCAAAAGCGCAGCAAAAGGCCAGGAACCGTAAAAAGCGCGGTTGCTGGCGTTTTTCCA 1920
|||
Db 1802 AGCAAAAGCGCAGCAAAAGGCCAGGAACCGTAAAAAGCGCGGTTGCTGGCGTTTTTCCA 1743
|||
QY 1921 TAGGCTCCGCCCCCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGCGAAA 1980
|||
Db 1742 TAGGCTCCGCCCCCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGCGAAA 1683
|||
QY 1981 CCCGACAGCACTATAAAGATACACAGGCGTTTCCCTCTGGAAGCTCCCTCGTGGCTCTCC 2040
|||
Db 1682 CCCGACAGCACTATAAAGATACACAGGCGTTTCCCTCTGGAAGCTCCCTCGTGGCTCTCC 1623
|||
QY 2041 TGTTCGCAACCTGCGCTTACCGGATACCTGTCCGCTTCTCCCTTCGGGAAGCGTGGC 2100
|||
Db 1622 TGTTCGCAACCTGCGCTTACCGGATACCTGTCCGCTTCTCCCTTCGGGAAGCGTGGC 1563
|||
QY 2101 GCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTTCGGTGTAGGTCTCGCTCCAAGCT 2160
|||
Db 1562 GCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTTCGGTGTAGGTCTCGCTCCAAGCT 1503
|||
QY 2161 GGGCTGTGTGACGAAACCCCGCTTACGCCGACCGCTGCGCTTATTCGGTTAACTATCG 2220
|||
Db 1502 GGGCTGTGTGACGAAACCCCGCTTACGCCGACCGCTGCGCTTATTCGGTTAACTATCG 1443
|||
QY 2221 TCTTGAGTCCAAACCGGTAAAGACACGCTTATCGCCACTTGGCAGCAGCACCTGTTAAACAG 2280
|||
Db 1442 TCTTGAGTCCAAACCGGTAAAGACACGCTTATCGCCACTTGGCAGCAGCACCTGTTAAACAG 1383
|||
QY 2281 GATTAGCAGAGCGAGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTTAACTA 2340
|||
Db 1382 GATTAGCAGAGCGAGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTTAACTA 1323
|||
QY 2341 CGGCTACACTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCAGCTTACCTTCGG 2400
|||
Db 1322 CGGCTACACTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCAGCTTACCTTCGG 1263
|||
QY 2401 AAAAAGCTTGGTAGCTCTTGATCCGCAAAACAAACACCGCTGTGTAGCGGTGGTTTTT 2460
|||
Db 1262 AAAAAGCTTGGTAGCTCTTGATCCGCAAAACAAACACCGCTGTGTAGCGGTGGTTTTT 1203
|||
QY 2461 TGTTCGAAGCAGCAGATTAACGCGCAGAAAAAAGGATCTCAAGAAGATCTCTTTGATCTT 2520
|||
Db 1202 TGTTCGAAGCAGCAGATTAACGCGCAGAAAAAAGGATCTCAAGAAGATCTCTTTGATCTT 1143
|||
QY 2521 TTCTACGGGCTGTGACGCTCAGTGGAACGAAACTCAAGTTAAGGATTTTGGTCATGAG 2580
|||
Db 1142 TTCTACGGGCTGTGACGCTCAGTGGAACGAAACTCACGTTAAGGATTTTGGTCATGAG 1083
|||
QY 2581 ATTATCAAAAAGGATCTTTCACCTAGATCTTTTAAATTTAAATTAAGTTTAAATCAAT 2640
|||
Db 1082 ATTATCAAAAAGGATCTTTCACCTAGATCTTTTAAATTTAAATTAAGTTTAAATCAAT 1023
|||
QY 2641 CTAAAGTATATATGAGTAAACTTGGTCTGACAGTTTACCAATGCTTAATCAGTGAGGACCC 2700
|||
Db 1022 CTAAAGTATATATGAGTAAACTTGGTCTGACAGTTTACCAATGCTTAATCAGTGAGGACCC 963
|||
QY 2701 TATCTCAGCGATCTGTCTATTTTGGTTTCAATGATGCTGCTGCTCCCGTCTGTAGAT 2760
|||
Db 962 TATCTCAGCGATCTGTCTATTTTGGTTTCAATGATGCTGCTGCTCCCGTCTGTAGAT 903
|||
QY 2761 AACTACGATACGGGAGGCTTACCATCTGGCCCCAGTGTGCAATGATACCCGAGACCC 2820
|||
Db 902 AACTACGATACGGGAGGCTTACCATCTGGCCCCAGTGTGCAATGATACCCGAGACCC 843
|||
QY 2821 ACGCTCACCGGCTCCAGATTTTATCAGCAATAAACACGACGAGCCGGAAGCGCGAG 2880
|||
Db 842 ACGCTCACCGGCTCCAGATTTTATCAGCAATAAACACGACGAGCCGGAAGCGCGAG 783
|||
QY 2881 AAGTGTCTCTGCACTTTATCGGCTTCCATCCAGTCTATTATTTGTTCCGCGGAAGCTAG 2940
|||

Db 782 AGTGGTCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTGTTGGCGGAGCTAG 723
Qy 2941 AGTAAGTAGTTCGCCAGTTAATAAGTTTGGCAACAGTTGTGTCATTTGCTACAGGCATCGT 3000
Db 722 AGTAAGTAGTTCGCCAGTTAATAAGTTTGGCAACAGTTGTGTCATTTGCTACAGGCATCGT 663
Qy 3001 GGTGTACAGCTCGTGGTGGTATGCTTCATTCAGTCCGGTCCCAACGATCAAGGCG 3060
Db 662 GGTGTACAGCTCGTGGTGGTATGCTTCATTCAGTCCGGTCCCAACGATCAAGGCG 603
Qy 3061 AGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTAGTCTCTTCGGTCTCCCGATCGT 3120
Db 602 AGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTAGTCTCTTCGGTCTCCCGATCGT 543
Qy 3121 TGTCAAGTAAGTTGGCGCAGTGTATCACTCATGGTTATGGCAGCACTGCAATAATTC 3180
Db 542 TGTCAAGTAAGTTGGCGCAGTGTATCACTCATGGTTATGGCAGCACTGCAATAATTC 483
Qy 3181 TCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTC 3240
Db 482 TCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTC 423
Qy 3241 ATTCTCAGAAATAGTGTATCGCGGACCGAGTTGCTCTTCGCCGGCGTCAATACGGGATAA 3300
Db 422 ATTCTCAGAAATAGTGTATCGCGGACCGAGTTGCTCTTCGCCGGCGTCAATACGGGATAA 363
Qy 3301 TACCGCGCCACATAGCAGAACTTTAAAGTGTCTCATCATTTGGAAAAAGTTCTTCGGGGCG 3360
Db 362 TACCGCGCCACATAGCAGAACTTTAAAGTGTCTCATCATTTGGAAAAAGTTCTTCGGGGCG 303
Qy 3361 AAAACTCTCAGGATCTTACCGTGTGTAGATCCAGTTCGATGTAAACCACCTCGTCACCC 3420
Db 302 AAAACTCTCAGGATCTTACCGTGTGTAGATCCAGTTCGATGTAAACCACCTCGTCACCC 243
Qy 3421 CAACTGATCTTCAGCATCTTTTACTTTTACAGCGCTTTTCTGGTGAGCAAAAAACAGGAAG 3480
Db 242 CAACTGATCTTCAGCATCTTTTACTTTTACAGCGCTTTTCTGGTGAGCAAAAAACAGGAAG 183
Qy 3481 GCAAAATCGCCAAAAAGGGAATAAGGCGCAGCAGGAAATGTTGAATCTCATACTCTT 3540
Db 182 GCAAAATCGCCAAAAAGGGAATAAGGCGCAGCAGGAAATGTTGAATCTCATACTCTT 123
Qy 3541 CTTTTTTCAATATTATTAAGCATTTTATCAGGGTTATTTCTCATGAGCGGATACATATT 3600
Db 122 CTTTTTTCAATATTATTAAGCATTTTATCAGGGTTATTTCTCATGAGCGGATACATATT 63
Qy 3601 TGAATGTATTTAGAAAAATAAACAATAGGGTTCCGCGCACATTTCCCGGAAAAAGTGCC 3660
Db 62 TGAATGTATTTAGAAAAATAAACAATAGGGTTCCGCGCACATTTCCCGGAAAAAGTGCC 3
Qy 3661 AC 3662
Db 2 AC 1

RESULT 10
LOCUS AX306330 4001 bp DNA linear PAT 11-DEC-2001
DEFINITION Sequence 13 from Patent WO0188121.
ACCESSION AX306330
VERSION AX306330.1 GI:17645569
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Plaetinck,G., Renard,J.P. and Bogaert,T.
TITILE Vector constructs
JOURNAL Patent: WO 0188121-A 13 22-NOV-2001;
Devgen NV (BE)
FEATURES Location/Qualifiers
source i. .4001
/organism="synthetic construct"

/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Plasmid pGN59A"
Query Match 71.3%; Score 2609.4; DB 6; Length 4001;
Best Local Similarity 83.1%; Pred. No. 0;
Matches 3089; Conservative 0; Mismatches 591; Indels 35; Gaps 7;
ORIGIN
Qy 1 CTAATTTGTAAGCGTTAATATTTTGTAAAAATTCGCGTTAAATTTTGTAAATCAAGTC 60
Db 63 CGAAATTTGTAAGCGTTAATATTTTGTAAAAATTCGCGTTAAATTTTGTAAATCAAGTC 122
Qy 61 ATTTTAAACCAATAGCGCGGAAATCGGCAAAATCCCTTATAATCAAAAGAATAGACCGA 120
Db 123 ATTTTAAACCAATAGCGCGGAAATCGGCAAAATCCCTTATAATCAAAAGAATAGACCGA 182
Qy 121 GATAGGTTGAGTGTGTTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGAACTGGACTC 180
Db 183 GATAGGTTGAGTGTGTTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGAACTGGACTC 242
Qy 181 CAACTCAAAAGGCGGCAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 240
Db 243 CAACTCAAAAGGCGGCAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 302
Qy 241 CTAATCAAGTTTGTGGGTTCGAGGTGCGCTTAAAGCACTAAATTCGGAACCTTAAAGGAG 300
Db 303 CAAATCAAGTTTGTGGGTTCGAGGTGCGCTTAAAGCTTAAATTCGGAACCTTAAAGGAG 362
Qy 301 CCCCCGATTTAGAGCTTGACGGGGAAGCCGCGCAACGTTGGCGAGAAAGGGAAGAA 360
Db 363 CCCCCGATTTAGAGCTTGACGGGGAAGCCGCGCAACGTTGGCGAGAAAGGGAAGAA 422
Qy 361 AGCGAAAGAGGCGGCGCTAGGCGCTGACAAAGTGTAGCGGTACGCTGCGGTAAACCAC 420
Db 423 AGCGAAAGAGGCGGCGCTAGGCGCTGACAAAGTGTAGCGGTACGCTGCGGTAAACCAC 482
Qy 421 CACACCGCGCGCTTAAATGCGCGCTACAGGGCGGTCCCATTCGCCATTCAGGCTGCG 480
Db 483 CACACCGCGCGCTTAAATGCGCGCTACAGGGCGGTCCCATTCGCCATTCAGGCTGCG 541
Qy 481 CAACTGTTGGGAAGGCGGATCGTGGGCGCTTTCGCTATTACGCCAGCTGCGCAAGG 540
Db 542 CAACTGTTGGGAAGGCGGATCGTGGGCGCTTTCGCTATTACGCCAGCTGCGCAAGG 601
Qy 541 GCGATGTCTGCAAGCGCATTAAGTTGGGTAAACGCCAGGGTTTCCAGTCACAGCTTG 600
Db 602 GCGATGTCTGCAAGCGCATTAAGTTGGGTAAACGCCAGGGTTTCCAGTCACAGCTTG 661
Qy 601 TAAAAACGAGCGCAGTGAGCGCGCT-----CGTTCAATTCAGCTTTTGAA 646
Db 662 TAAAAACGAGCGCAGTGAATTTGTAATACGACTCACTATAGGGCGAATTCGAGCTCGGTAC 721
Qy 647 CCGGTGGAGGAGCGGCAGACTCGCGGTGCAATGTGTTTACAGCGTGTAGGAGAGATG 706
Db 722 CCGGGGATCTCTAGAGATCCCTCGACCTCGAGATCCATTGTGCTGGAAGGATCTGGAT 781
Qy 707 AAGATGCTCGACACGCTGCGAGAACACGACAGCTAGAT-----TAACCTCTAGAAAGATAA 759
Db 782 CCGGCTTACTAAAGCCAGATAACAGTATGCGTATTTGCGCGCTGATTTTTCGGGTATAA 841
Qy 760 TCATATTTGACGTACGTTTAAAGATAATCATCGTAAAAATG-----ACGATCGGAT 812
Db 842 GAAATATATCTGATATGTATACCCGAGATGTGTCAAAAAGAGGTGTCTATGAAGCAGCG 901
Qy 813 CTGTAATACGACTCACTATAGGCGGAAATGGGTACCGGGCCCCCTCGAGGTCCGCGT 872
Db 902 TATTACGTGACAGTTTGACAGCGACAGCTATCAGTTGCTCAAGGGCATATATGATGTCAAT 961
Qy 873 ATCGATAAGCTTTGATATCGAATTCCTGACGCGCGGGGATCCACTAGTTCTTAGAGCGGCC 932
Db 962 ATCTCGGTCTCGTAAGCAACCAATGCAGATGAAGCCCGTCTGCTGCTGCCGACGC 1021

Db	3181	AAAGCGGTTAGCTCCTTCGGTCCGATCGTGTGTCAGAGTAAGTTGGCCGACGTGT	3240
QY	3148	ATCAGTCATGGTTATGCGAGCACTGCATAAATCTCTTACATGTCATGCCATCCGTAAGATG	3207
Db	3241	ATCAGTCATGGTTATGCGAGCACTGCATAAATCTCTTACATGTCATGCCATCCGTAAGATG	3300
QY	3208	CTTTTCGTGACTGTGGTGAAGTACTCAACCAAGTCAITTCGAGATAGTATGCGGCGACC	3267
Db	3301	CTTTTCGTGACTGTGGTGAAGTACTCAACCAAGTCAITTCGAGATAGTATGCGGCGACC	3360
QY	3268	GAGTTGCTCTTTCGCGCGCTCAATACGGGATATACCGGCCACATAGCAGAACTTTAAA	3327
Db	3361	GAGTTGCTCTTTCGCGCGCTCAATACGGGATATAGTATGATGATAGCAGAACTTTAAA	3420
QY	3328	AGTGCTCATATTGGAAGAGTTCTTCGGGGCGAAGCACTCTCAAGATCTTACCGCTGT	3387
Db	3421	AGTGCTCATATTGGAAGAGTTCTTCGGGGCGAAGCACTCTCAAGATCTTACCGCTGT	3480
QY	3388	GAGATCCAGTTCGATTAACCACTCGTCGCCACCACTGATCTTCAGCATCTTTTACTTT	3447
Db	3481	GAGATCCAGTTCGATTAACCACTCGTCGCCACCACTGATCTTCAGCATCTTTTACTTT	3540
QY	3448	CACCAGCGTTTCTGGTGAGCAAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAG	3507
Db	3541	CACCAGCGTTTCTGGTGAGCAAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAG	3600
QY	3508	GCGCACCGGAATGTTGAATACTCATCTCTTCTTTTCAATATATGAGCAATTA	3567
Db	3601	GCGCACCGGAATGTTGAATACTCATCTCTTCTTTTCAATATATGAGCAATTA	3660
QY	3568	TCAGGGTTATCTCATGAGCGGATACATATTTGAAATGATTTAGAAAAATACAAAT	3627
Db	3661	TCAGGGTTATGTTCTCATGAGCGGATACATATTTGAAATGATTTAGAAAAATACAAAT	3720
QY	3628	AGGGGTTCCGCGCACATTTTCCCGGAAAGTGCCAC	3662
Db	3721	AGGGGTTCCGCGCACATTTTCCCGGAAAGTGCCAC	3755
RESULT 11			
AF179627			
LOCUS	AF179627	5540 bp	DNA linear SYN 22-NOV-1999
DEFINITION	Expression vector pCENEX645, complete sequence.		
ACCESSION	AF179627		
VERSION	AF179627.1	GI:6457303	
KEYWORDS	Expression vector pCENEX645		
SOURCE	Expression vector pCENEX645		
ORGANISM	Other sequences; artificial sequences; vectors.		
REFERENCE	1 (bases 1 to 5540)		
AUTHORS	Eugster,H.P., Bartsch,S., Murgler,F.E. and Sengstag,C.		
TITLE	Functional co-expression of human oxidoreductase and cytochrome P450 1A1 in Saccharomyces cerevisiae results in increased EROD activity		
JOURNAL	Biochem. Biophys. Res. Commun. 185 (2), 641-647 (1992)		
MEDLINE	92304288		
PUBMED	1610357		
REFERENCE	2 (bases 1 to 5540)		
AUTHORS	Sengstag,C. and Paladino,G.		
TITLE	A collection of transformed Saccharomyces cerevisiae strains as an unlimited source of human enzymes		
JOURNAL	BioWorld 97 (6), 2-7 (1997)		
REFERENCE	3 (bases 1 to 5540)		
AUTHORS	Sengstag,C.		
TITLE	Direct Submission		
JOURNAL	Submitted (24-AUG-1999) Genetics, Institute of Toxicology, ETH Zurich, Schorenstrasse 16, Schwerzenbach 8603, Switzerland		
FEATURES	Location/Qualifiers		
source	1..5540		
	/organism="Expression vector pCENEX645"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:106613"		
	/note="Yeast expression vector pCENEX645 (alias pC9645)"		

		containing truncated GAPDH promoter for constitutive expression of CDNA _s in yeast"	
		417..1220	/gene="URA3"
		417..1220	/note="URA3"
		417..1220	/note="yeast selection marker"
		/codon_start=1	/product="URA3"
		/protein_id="AAF09466.1"	/db_xref="GI:6457304"
		/translation="MSKATYKERAATHPSVAAKLFNIMHEKQTNLCASLDVRRTKEL	
		LELPAIGPKICLLATHVDILTFPSMEGTVPKPKALSAKYNFLPDRFADIGNTVK	
		LOYSAGVTRIAEWADITNHNHGVGIGVSLGKQAAEEVTKPRGLLMLAELCKGSLIS	
		TGEYTKGTVDIAKSDKDFVIGTIAQRDMGRDEGYDMLTWPVGVLDDIGDALGQQYR	
		TVDDVVSFGSDIIVIGRLFAKGRDAKVEGERYRKAGWEAYLRRCQQN"	
		2310..2508	/note="truncated GAPDH promoter"
		complement(3958..4818)	/gene="Amp"
		complement(3958..4818)	/gene="Amp"
		marker="ampicillin resistance; Escherichia coli selection	
		/codon_start=1	/product="Amp"
		/protein_id="AAF09467.1"	/db_xref="GI:6457305"
		/translation="MSIOHFRVALIPFFAAFLPFAHPDLPLVFKVDCDQLGARVG	
		IELDLSNGKILESPEERPEPMWSTFKVLLCGAVLFRIDAGOBQLGRRTHYSNDLVE	
		YSPVTEKHLTDGMTYRELCSAITSNDNTAANLLTIGGPKELTAFLEHMDGHVTRL	
		DRWEPLENAIPNDRDITMPVAMATTLKLLTGBELLTLASRQQLIDWNEADKVGAPL	
		LRSLAUPFATADKSGAGERSGRGIIAALGPDGKPSRIVVIYTTGSGQATMDERNQIA	
		EIGASLIKHW"	
		ORIGIN	
		Query Match	69.9%; Score 2558.8; DB 12; Length 5540;
		Best Local Similarity	82.8%; Pred. No. 0;
		Matches 3035; Conservative	0; Mismatches 557; Indels 72; Gaps 7;
QY	3	AAATTTGTAAGCGTTAAATATTTTGTAAATTCGCGTAAATTTTGTAAATCAGCTCAT	62
Db	1351	AAATTTGTAAGCGTTAAATATTTTGTAAATTCGCGTAAATTTTGTAAATCAGCTCAT	1410
QY	63	TTTTTAACCAATAGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAAATAGACCGAGA	122
Db	1411	TTTTTAACCAATAGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAAATAGACCGAGA	1470
QY	123	TAGGGTTGAGTGTGTTTCCAGTTTGGAAACAAGATCCACTATTAAGAAGCGTGGACTCCA	182
Db	1471	TAGGGTTGAGTGTGTTTCCAGTTTGGAAACAAGATCCACTATTAAGAAGCGTGGACTCCA	1530
QY	183	ACGTCAAAAGGGCGAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCT	242
Db	1531	ACGTCAAAAGGGCGAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCT	1590
QY	243	AATCAAGTTTTTGGGGTTCAGGTCCGTAAGCACTAAATCGGAACCTCTAAAGGAGGCC	302
Db	1591	AATCAAGTTTTTGGGGTTCAGGTCCGTAAGCACTAAATCGGAACCTCTAAAGGAGGCC	1650
QY	303	CCGATTTAGACTTTCAGGGGAAAGCCCGCAAGTGGCGGAGAAAGGAGGAGGAGG	362
Db	1651	CCGATTTAGACTTTCAGGGGAAAGCCCGCAAGTGGCGGAGAAAGGAGGAGGAGG	1710
QY	363	CGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTACGCTGCCGTAAACCA	422
Db	1711	CGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTACGCTGCCGTAAACCA	1770
QY	423	CACCCGCGCGCTTAATGCGCGCTACAGGGCGCGTC---CCATTGCGCATTCAGGCTGC	479
Db	1771	CACCCGCGCGCTTAATGCGCGCTACAGGGCGCGTC---CCATTGCGCATTCAGGCTGC	1830
QY	480	GCAACTGTTGGGAAGGGCGATCGGTGCGGGCCCTCTTCGCTATTACGCCAGCTGGCGG	539

1831 GCACCTGTTGGAAAGGGCGATCGGTGGGGCTCTTGGCTATTACCGCAGCTGGCGAAAG 1890
540 GGGGATGTCTCGAAGGCGATTAAAGTTGGTTAAAGCCAGGGTTTCCAGTCAAGACGTT 599
1891 GGGGATGTCTCGAAGGCGATTAAAGTTGGTTAAAGCCAGGGTTTCCAGTCAAGACGTT 1950
600 GTAAACGAGCGCCAGTGAAGCGCGCTCGTTCAATTCACGTTTTTGAACCGGTGGAGGAG 659
1951 GTAAACGAGCGCCAGTGA-----ATTGTAATACGACTCACTATAGGCGGAATTGG 2001
660 GGCAGACTCGCGGTGCAATGTGTTTTACAGCGTGAATGAGCAGATGAAGATGCTCGACA 719
2002 AGCTCCACCGCGTG-----GGGCGGCTCTAGTCGACGCTCTGCC 2042
720 CCTGTCAGAACACGAGCTAGATTAAACCTAGAAAGATTAATCATATTGTGAGCTAGCTTA 779
2043 TTATCGGACTCTGCTAGTTAGGAAGCAGCCAGTAGTAGTTGAGGCGGTTGAGCACCGCC 2102
780 AAGATAATCATGCGTAAATTCACGCATGGGATCTGTAATACGACTCACTATAGGCGGAA 839
2103 GCCGCAAGGAATGTGTCATGAAGGAGATGGCGCCCAACAGTCCCGGCCACGGGCGCT 2162
840 TTGGGTACCGGCGCCCGCTCGAGTTCGACGGTATCGATAGCTTTGATATCGAATTCCTG 899
2163 GCCACCATACCCACGCCGAAACAAAGCGCTCATGAGCCGGAAGTGGCGAGCCCGATCTTCC 2222
900 CAGCCCGGGGATCCACTAGTTCTAGAGCGGCGGCCACCGCGGTGGAGCTCAGCTTTTG 959
2223 CCATCGGTGATGTCGGCGATATAGGCGGCGAGCAACCGCACCTGTGGCGCGGTGATGCG 2282
960 TTCCCTTTAGTGAGGTTAAATPAGATCCATCGCTCAATTTTACGCACTATCTTTCTA 1019
2283 GCCACGATGCTCGCGGTAGGATCTCCCTGAAAGGTTGAAACAGTTCCTG 2342
1020 GGGTTAACTAGCTGCATCAGGATCATATCGTGGGTCTTTTTCGGCTCAGTCAATCGC 1079
2343 AAATTTATC-----CCCTACTTGACTAATAAGTATAAA 2376
1080 CCAAGCTGGCGTATCTGGGCAATCGGGAGGAAGCCCGTCTTTTCCCGCGAGTT 1139
2377 AGACGGTAGGTATGATTGTAAATCTGTA-----AATCTATTCTTAAACTTCTTAAAT 2431
1140 GAAGCGGATGAAAGAGTTTCGCGAGGATGACTGCTGCTGATTAAGCTTGAGCGGAAAA 1199
2432 CTACTTTATAGTTAGTCTTTTATTTAGTT-----TTAAACACCAAGAACTTAGT 2482
1200 CGCAGTTTACCATGATTCGGGAAGGTGGCCATGCAAGCTTTTAAACGGTGAACCTG 1259
2483 TTGGAATAAACACACATAAATAAGAAATTCATGCTTTTGCAGCTTTCTTTTCTCTTTTG 2542
1260 TTCGTTTCAGGCACCTGGGATACCAAGTTCGTGCGGCTTTTCCGACACAGTTCCGGATG 1319
2543 GCTGTTTTGACGCAAAATATCTGCATCTTAGCTCTCGAGGGGAACACTACTCATTAAC 2602
1320 GTCAGCCGGAAGCGCATCAGCAACCGAAACATAACCGGCAAGCGGAACTGCGCTGCC 1379
2603 AACGCCAGTCTATGAGACAAATAGTTTGTATACTAATAATATGGAACATAATAGC 2662
1380 GGTGTGCAAGTTAATGACAGCGGTGGCGCTGGGATATATAGTCAGCGAGGACGGGTAT 1439
2663 AATACCCAAATTTTATCTAAATTTGCGGAAAGATTAATAATCAGCTTATCGATAACCGT 2722
1440 CTTGCTGATCCCGCAGAAATGGACATGGATACCCCGTAGTTACCGG-GCGGGCGGC 1498
2723 CAGGGGGGCGCGCGCGCCCGCTTTGTTCCCTTTAGTGGGGTTAATCCGAGC 2782
1499 TTGGCGTAATCATGTGTCATAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCAACAATCCA 1558
2783 TTGGCGTAATCATGTGTCATAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCAACAATCCA 2842
1559 CACAAATACGAGCGGGAAGCATAAAGTGTAAAGCTGGGGTGCCTTAATGAGTGAAGCTAA 1618

2843 CACAAATACGAGCGGGAAGCATAAAGTGTAAAGCTGGGGTGCCTTAATGAGTGAAGTTAA 2902
1619 CTCACATTAATTTGCGTTGCGTCACTGCGCGCTTTCCAGTCGGGAAACCTGTCGTGCCAG 1678
2903 CTCACATTAATTTGCGTTGCGTCACTGCGCGCTTTCCAGTCGGGAAACCTGTCGTGCCAG 2962
1679 CTGCAATTAATCGGCAACGCGCGGGGAGAGCGGTTTCGTAATTGGGCGCTCTTCC 1738
2963 CTGCAATTAATCGGCAACGCGCGGGGAGAGCGGTTTCGTAATTGGGCGCTCTTCC 3022
1739 GCTTCTCGCTCACTGACTCGCTGCGCTCGTTCGCTCGGCGGAGCGGTATCAGCT 1798
3023 GCTTCTCGCTCACTGACTCGCTGCGCTCGTTCGCTCGGCGGAGCGGTATCAGCT 3082
1799 CACTCAAAAGGGGTAAATACGGTTATCCACAGATCAGGGATAAACGAGGAAGAACATG 1858
3083 CACTCAAAAGGGGTAAATACGGTTATCCACAGATCAGGGATTAACGAGGAAGAACATG 3142
1859 TGAGCAAAAGGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCTTGCTGGCGCTTTTC 1918
3143 TGAGCAAAAGGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCTTGCTGGCGCTTTTC 3202
1919 CATAGGCTCGCGCCCGCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGTGGCGCA 1978
3203 CATAGGCTCGCGCCCGCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGTGGCGCA 3262
1979 AACCCGACAGGACTATAAGATACAGCGGTTTCCCGCTGGAAGCTCCCTCGTGGCTCT 2038
3263 AACCCGACAGGACTATAAGATACAGCGGTTTCCCGCTGGAAGCTCCCTCGTGGCTCT 3322
2039 CCTGTTCCGACCGCTCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTG 2098
3323 CCTGTTCCGACCGCTCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTG 3382
2099 GCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTGGGTAGGTGCTCGCTCCAG 2158
3383 GCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTGGGTAGGTGCTCGCTCCAG 3442
2159 CTGGGCTGTGTGACGAAACCCCGCTCAGCCGACCGCTGCGCTTATCCCGTAACTAT 2218
3443 CTGGGCTGTGTGACGAAACCCCGCTCAGCCGACCGCTGCGCTTATCCCGTAACTAT 3502
2219 CGTCTTGTAGTCAACCCCGTGAACACGACTTATCGCCACTGGCAGAGCCACTGGTAAC 2278
3503 CGTCTTGTAGTCAACCCCGTGAACACGACTTATCGCCACTGGCAGAGCCACTGGTAAC 3562
2279 AGGATTAGCAGAGCAGGTATGTAGCGGTGCTACAGAGTTCTTCAAGTGGTGGCTAAC 2338
3563 AGGATTAGCAGAGCAGGTATGTAGCGGTGCTACAGAGTTCTTCAAGTGGTGGCTAAC 3622
2339 TACGGCTACACTAGAAAGGACAGTATTTGGTATCTCGGCTCTGCTGAAGCCAGTTACCTTC 2398
3623 TACGGCTACACTAGAAAGGACAGTATTTGGTATCTCGGCTCTGCTGAAGCCAGTTACCTTC 3682
2399 GGAATAAGAGTTGTAGCTCTTGATCCGCAAAACCAACCGCTGGTAGCGGTGGTTTT 2458
3683 GGAATAAGAGTTGTAGCTCTTGATCCGCAAAACCAACCGCTGGTAGCGGTGGTTTT 3742
2459 TTTGTTTGCAGCAGCAGATTACCGCGCAGAAAAAAGGATCTCAAGAGATCCTTTGATC 2518
3743 TTTGTTTGCAGCAGCAGATTACCGCGCAGAAAAAAGGATCTCAAGAGATCCTTTGATC 3802
2519 TTTTCTCAGGGGTCTGAGCTCAGTGGAAACGAAAACTCAGCTTAAGGGATTTTGTGATG 2578
3803 TTTTCTCAGGGGTCTGAGCTCAGTGGAAACGAAAACTCAGCTTAAGGGATTTTGTGATG 3862
2579 AGATTATCAAAAGGATCTTCACTAGATCTTTTAAATTTAAATTAAGATTTTAAATCA 2638
3863 AGATTATCAAAAGGATCTTCACTAGATCTTTTAAATTTAAATTAAGATTTTAAATCA 3922
2639 ATCTAAGTATATAGTAACTTTGCTGACAGTTACCAATGCTTAATCAGTGGGCA 2698
3923 ATCTAAGTATATAGTAACTTTGCTGACAGTTACCAATGCTTAATCAGTGGGCA 3982

QY	2699	CTATCTCAGCGATCTGTCTATTCTTCTTCATTCATATAGTTGCTCTGACTCCCGTCTGRTAG	2758
Db	3983	CTATCTCAGCGATCTGTCTATTCTTCTTCATTCATATAGTTGCTCTGACTCCCGTCTGRTAG	4042
QY	2759	ATAACTACCATACGGGAGGCTTACCATTCTGGCCCAAGTCTGCAATGATACCGCGAGAC	2818
Db	4043	ATAACTACCATACGGGAGGCTTACCATTCTGGCCCAAGTCTGCAATGATACCGCGAGAC	4102
QY	2819	CCACGCTCACCGCTCCAGATTATCAGCAATAAACACGACCGCGGAAGGCGCGAGCGC	2878
Db	4103	CCACGCTCACCGCTCCAGATTATCAGCAATAAACACGACCGCGGAAGGCGCGAGCGC	4162
QY	2879	AGAAGTGTCTCGCAACTTTATCCGCTCCATCAGTCTATTAAATTTGTTGCGGGAAGCT	2938
Db	4163	AGAAGTGTCTCGCAACTTTATCCGCTCCATCAGTCTATTAAATTTGTTGCGGGAAGCT	4222
QY	2939	AGAGTAAGTAGTTCCGCACTTAATAGTTGCGCAAGTGTGGCCATTGCTACAGGCATC	2998
Db	4223	AGAGTAAGTAGTTCCGCACTTAATAGTTGCGCAAGTGTGGCCATTGCTACAGGCATC	4282
QY	2999	GTGGTGTACGCTCGTGTGTTGATGGCTTCATTGAGTCCGGTTCCTCAACGATCAAGG	3058
Db	4283	GTGGTGTACGCTCGTGTGTTGATGGCTTCATTGAGTCCGGTTCCTCAACGATCAAGG	4342
QY	3059	CGAGTTACATGATCCCGCATGTTGTCAGCAAAAAGCGGTGAGTCTCTTCGGTCTCCGATC	3118
Db	4343	CGAGTTACATGATCCCGCATGTTGTCAGCAAAAAGCGGTGAGTCTCTTCGGTCTCCGATC	4402
QY	3119	GTGTGAGAGTAGTTGCGCGCAGTGTATCACTCATGTTGATGCGCAGCTGCATAT	3178
Db	4403	GTGTGAGAGTAGTTGCGCGCAGTGTATCACTCATGTTGATGCGCAGCTGCATAT	4462
QY	3179	TCTTCTACTGTATGTCATCCGTAAGTCTTCTGCTGACTGGTGGTACTCAACCAAG	3238
Db	4463	TCTTCTACTGTATGTCATCCGTAAGTCTTCTGCTGACTGGTGGTACTCAACCAAG	4522
QY	3239	TCATTCTGAGATAGTGTATGCGGCACCGAGTGTCTTTCGCGCGCTCAATACGGGAT	3298
Db	4523	TCATTCTGAGATAGTGTATGCGGCACCGAGTGTCTTTCGCGCGCTCAATACGGGAT	4582
QY	3299	ATACCGCGCCACATGACGAGACTTTAAAGTCTCATCTTGGAAAGCTTCTTCGGG	3358
Db	4583	ATACCGCGCCACATGACGAGACTTTAAAGTCTCATCTTGGAAAGCTTCTTCGGG	4642
QY	3359	CGAAACTCTCAAGGATCTTACCGCTGTGAGATCCAGTTCGATGTAACCCACTCGTGCA	3418
Db	4643	CGAAACTCTCAAGGATCTTACCGCTGTGAGATCCAGTTCGATGTAACCCACTCGTGCA	4702
QY	3419	CCCACTGATCTTCAGCATCTTTTACTTTTCCAGCGCTTCTGGGTGACGCAAAAACAGGA	3478
Db	4703	CCCACTGATCTTCAGCATCTTTTACTTTTCCAGCGCTTCTGGGTGACGCAAAAACAGGA	4762
QY	3479	AGGCAAAATGCGCGCAAAAAGGAATAAGGCGGACAGCGGAATGTTGAATCTCATCTC	3538
Db	4763	AGGCAAAATGCGCGCAAAAAGGAATAAGGCGGACAGCGGAATGTTGAATCTCATCTC	4822
QY	3539	TTCCCTTTTCAATATTATTGAGCATTTTATCAGGGTTATTTGCTCATGAGCGGATACATA	3598
Db	4823	TTCCCTTTTCAATATTATTGAGCATTTTATCAGGGTTATTTGCTCATGAGCGGATACATA	4882
QY	3599	TTTGAATGTATTAGAAAAATAAACAATAGGGGTTCCGCGCACAATTTCCCGGAAAAATG	3658
Db	4883	TTTGAATGTATTAGAAAAATAAACAATAGGGGTTCCGCGCACAATTTCCCGGAAAAATG	4942
QY	3659	CCAC 3662	
Db	4943	CCAC 4946	
RESULT 12			
ASPGREEN2/c			
LOCUS			
3633 bp DNA linear SYN 15-MAY-1998			

DEFINITION	Artificial sequences, plasmid vector pGreen-2.
ACCESSION	Y09374
VERSION	Y09374.1 GI:1684627
KEYWORDS	blatEM gene; lacZ-PhoC fusion gene; multiple cloning site; T3 promoter; T7 promoter.
SOURCE	synthetic construct
ORGANISM	synthetic construct
REFERENCE	1
AUTHORS	Thaller, M.C., Berlutti, F., Schippa, S., Selan, L. and Rossolini, G.M.
TITLE	Bacterial acid phosphatase gene fusions useful as targets for cloning-dependent insertional inactivation
JOURNAL	Biotechnol. Prog. 14 (2), 241-247 (1998)
MEDLINE	98215688
PUBMED	9548775
REFERENCE	2 (bases 1 to 3633)
AUTHORS	Rossolini, G.M.
TITLE	Direct Submision
JOURNAL	Submitted (11-NOV-1996) G.M. Rossolini, Dip. Biologia
COMMENT	Molecular-Seq., Microbiologia,, Univ. di Siena, via Laterina N.8, I- 53100 Siena, ITALY
FEATURES	Related sequences: X6444 and X52328.
source	1..3633
	Location/Qualifiers
	/organism="synthetic construct"
	/mol_type="other DNA"
	/db_xref="taxon:32630"
	/plasmid="pGreen-2"
	129..989
gene	/gene="blatEM"
CDS	129..989
	/gene="blatEM"
	/codon_start=1
	/transl_table=11
	/protein_id="CAA70549.1"
	/db_xref="GI:1684628"
	/translation="MSIQHFRVALIPFFAFCLPVFAHPETLVKVDADQLGARVGY
	IELDLNSGKILDESFPREPFPMWTFKLLGVLRDQWENPGVTHGAIIPAGNDATTPDLYYL
	YSPVTEKHLDTGMTRVELCSAAITMSDNTAANLLLTIGGPKELTAFLHMGDVTPL
	DRWPELNEATPNDRDRTMPVMATTLRLKLLTGELLTLASRQLIDMNEADKVAQPL
	LRSLALPAGWFTADKSGAGERSGRIIAALGPDGKPSRIIVIVTTGSGQATMDERNQIA
	ETGASLIQKH"
primer_bind	2147..2162
	/note="PUC M13/Reverse Sequencing primer recognition site"
gene	2154..3104
	/gene="lacZ-PhoC"
CDS	2154..3104
	/gene="lacZ-PhoC"
	/codon_start=1
	/transl_table=11
	/protein_id="CAA70550.1"
	/db_xref="GI:1684629"
	/translation="MTMITPSAQLTLTKGNKSWSTAVAAALELVDPGCCRNSSISLS
	IPSTRGPGVPNSPYSESYARSIAVLQRRDQWENPGVTHGAIIPAGNDATTPDLYYL
	KNEAIDSLKLPPEVSGISQFLNDQAMYERMLRNTERKQQAQADLAAGVAT
	AFSGAFGPTTEKSPFLYKLTNNIEDAGLATRSKAEHYMRIIPFAYGTCTNK
	DKDLSTNGSPSGHTSIGWATLVLAEPVNPANQAILERGYQLGQSRVICGTHWQSD
	VDAARIVGSAVAATLHSDPAFOAQLAKAKQBEFAKQSKDL"
promoter	2179..2198
	/gene="lacZ-PhoC"
	/note="T3 promoter"
misc_feature	2214..2316
	/gene="lacZ-PhoC"
promoter	/note="multiple cloning site"
	2325..2344
	/gene="lacZ-PhoC"
primer_bind	/note="T7 promoter"
	2354..2370
	/gene="lacZ-PhoC"
primer_bind	/note="M13-20 sequencing primer recognition site"
	2374..2390
	/gene="lacZ-PhoC"
	/note="M13-40 sequencing primer recognition site"

ORIGIN	Query Match 69.7%; Score 2550.6; DB 12; Length 3633; Best Local Similarity 82.3%; Pred. No. 0; Matches 3020; Conservative 0; Mismatches 604; Indels 47; Gaps 6;		
Qy	1 CTAAATTGTAAGCGTTAAATATTTTGTAAATTTTCGCGTTAAATTTTGTAAATTCAGCTC 60	Qy	1021 GGTTAATCTAGCTGCATCAGGATCATATCGTGGGTCCTTTTTCGGGCTCAGTCACTCGCC 108
Db	3633 CTAAATTGTAAGCGTTAAATATTTTGTAAATTTTCGCGTTAAATTTTGTAAATTCAGCTC 3574	Db	2627 CACCCCGCTTGGCGGCAGGTTCAGCATCTGCTGTGCTGTTTTCGGCGCTCGGATTTGG 2568
Qy	61 ATTTTAAACCAATAGGCGGAATTCGCAAAATCCCTTATAAATCAAAAGAAATAGACCGA 120	Qy	1081 CRAGCTGGGCTATCTGGGCATCGGGAGAGAAAGCCCGTGCCTTTTCCCGCGAGGTTG 1140
Db	3573 ATTTTAAACCAATAGGCGGAATTCGCAAAATCCCTTATAAATCAAAAGAAATAGACCGA 3514	Db	2567 CAGCATACGCCCTTCTCATACATTCCTGTATCATTTTAAATCTGAATATCTGCCGACTTC 2508
Qy	121 GATAGGTTGAGTGTGTTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGAACGTGGACTC 180	Qy	1141 AAGCGG-CATGGAAGAGTTTGCAGGAGTGAATCTGCTGCTGCAATTCAGCTTGAGCGAAA 1199
Db	3513 GATAGGTTGAGTGTGTTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGAACGTGGACTC 3454	Db	2507 CGCGGTTGGCGGTAAACAGITTCAGGCTGTCGATAGCTGTTCATTTTTCAGATAATATA 2448
Qy	181 CAAGCTCAAAAGGGGAAAAACCGTCTATCAGGGCGATGGCCACTAGTGAAACCATCACC 240	Qy	1200 CGCACGTTTACCATGATGATTCGGGAAGGTGTGGCCATGCAAGCCCTTTAAACGTTGAAC 1259
Db	3453 CAAGCTCAAAAGGGGAAAAACCGTCTATCAGGGCGATGGCCACTAGTGAAACCATCACC 3394	Db	2447 ATCCGGCTTGGTGGCATCGTTGCTTCCCGGGATCGCTCCATCGGTTAAACGCCAGGGTT 2388
Qy	241 CTAATCAAGTTTTTTTGGGTCGAGTCCGCTTAAGACACTAAATCGGAACCTTAAGGGAG 300	Qy	1260 TTCTGTTTCAGGCCACCTTCGGGATACCAAGTTCTGTCGCGCTTTTCCGGACACAGTTTCCG 1319
Db	3393 CTAATCAAGTTTTTTTGGGTCGAGTCCGCTTAAGACACTAAATCGGAACCTTAAGGGAG 3334	Db	2387 TTCCCGAGTCACGACGTTGTAAACACGACCGGCAGTGTAGCGCGCTTAATACGATCACTATA 2328
Qy	301 CCCCGATTTAGACTTTCAGCGGGAAGCGCGGAACGTCGGCGAGAAAGGAGGAGAA 360	Qy	1320 GTCAAGCCGAAGCGCATCAGCAACCCGAAACAATACCGGGGACAGCGCGGATTAATTCGCGCT 1379
Db	3333 CCCCGATTTAGACTTTCAGCGGGAAGCGCGGAACGTCGGCGAGAAAGGAGGAGAA 3274	Db	2327 G-----GGCGAATTCGGTACCGGGCCCCCTCGAGTCTGACGATCGATTAAGCT 2277
Qy	361 AGCGAAAGGAGCGGCGCTAGGCGCTGGCAAGTGTAGCGGTCAAGTTCGCGGTAAACAC 420	Qy	1380 GGTGTCCAGATTAAATGACAGCGGTTCGGCGCTGGGATATTAATCTGACGAGGAGCGGTTAT 1439
Db	3273 AGCGAAAGGAGCGGCGCTAGGCGCTGGCAAGTGTAGCGGTCAAGTTCGCGGTAAACAC 3214	Db	2276 TGATATCGAATTCCTGCAGCCGGGGATCCACTAGTTCTTAGAGCGCGCGCCACCGCGGT 2217
Qy	421 CACACCGCGCGCTTAATGCGCGCTACAGGGCGCTCCATTCGCGCATTCAGGCTGCG 480	Qy	1440 CCTGGCTGATCGCCGAGAAATGGACATGGATACCCGCTGAGTTTACCGCGCGCGCGCT 1499
Db	3213 CACACCGCGCGCTTAATGCGCGCTACAGGGCGCTCCATTCGCGCATTCAGGCTGCG 3154	Db	2216 GGAGCT-----CCAGCTTTTGTTCCTTTAGTAGAGGTTAAATTCGCGCT 2172
Qy	481 CAACTGTTGGGAAGGGCGATCGGTGCGGCTCTTCGCTATTAGCCGAGCTGGCGAAAG 540	Qy	1500 TGCGGTAATCATGCTCATAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCACAAATTCAC 1559
Db	3153 CAACTGTTGGGAAGGGCGATCGGTGCGGCTCTTCGCTATTAGCCGAGCTTACAGATCTT 3094	Db	2171 TGCGGTAATCATGCTCATAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCACAAATTCAC 2112
Qy	541 GGGATGTCTGCAAGGCGATTAAGTTGGGTAACGCCAGGCTTTTCCAGTTCACGAGTTG 600	Qy	1560 ACAACATACGAGCCGGAAGCATAAAGTGTAAAGCTTGGGTTGCTTAATGAGTGAGTTAAC 1619
Db	3093 TCTGTGATTTTGTGCAAAATCTCTGTTTGGCTTTCGCTAATGCGCTGAAATGCGGAT 3034	Db	2111 ACAACATACGAGCCGGAAGCATAAAGTGTAAAGCTTGGGTTGCTTAATGAGTGAGTTAAC 2052
Qy	601 TAAACACGCGCGAGTGAAGCGCGCTCGTTCAATTCAGCTTTTGTAAACCGTGGAGAGG 660	Qy	1620 TCACATTAATTTGGTTGCGCTCATCTGCGCGCTTTCAGTTCGGGAACCTGTCTGCTCCAG- 1678
Db	3033 CGGAATGAATGTGCGACAGCGGCTGAACCGCAATTCGCGCGGCAATCCACATCACTCT 2374	Db	2051 TCACATTAATTTGGTTGCGCTCATCTGCGCGCTTTCAGTTCGGGAACCTGTCTGCTCCAGC 1992
Qy	661 GCAGACTCGGCTCAAAATGTTTACAGCTGATGGAGCAGATGAAGATGCTCGACAC 720	Qy	1679 -----CTGCATTAATGAATCGCCAAACGCGCGGGAGAGGCGGTTTGGTATTCGGCG 1731
Db	2973 GCCAGTGAATGCGCAAAATCACCC-----GGCTCTGTCGAGCTGATAACCCC 2926	Db	1991 TGCATCCTTCGATTAATGAATCGGCCAAACGCGCGGGAGAGGCGGTTTGGTATTCGGCG 1932
Qy	721 GCTGCGAACAACGAGCTAGATTAAACCTTAGAAGATAAATCATATTGTGACGTACGTTAA 780	Qy	1732 CTCTTCGCTTCTCTGCTCACTGACTCTGCTGCGCTCGTTGCGCTCGGCGAGCGGT 1791
Db	2925 GTTCCAGAAATCTGATTTTCCCGGTTTCACTTCGCGCAGCACTAATGCGGTTGCGCC 2866	Db	1931 CTCTTCGCTTCTCTGCTCACTGACTCTGCTGCGCTCGTTGCGCTCGGCGAGCGGT 1872
Qy	781 AGATAATCATGCGTAAATTTGACGATGGGATCTGTAAATACGACTCACTATAGGCGGAAT 840	Qy	1792 ATCAGCTCACTCAAGGGCGGTAATACGGTTATCCACAGAAATCAGGGGATAACACGAGAAA 1851
Db	2865 AGCCGATAGACGTATGACCTCGAGGTAAGATTCGTTTGGTGGAGAGTTTTCCTGATCTT 2806	Db	1871 ATCAGCTCACTCAAGGGCGGTAATACGGTTATCCACAGAAATCAGGGGATAACACGAGAAA 1812
Qy	841 TGGGTACCGGGCCCCCTCGAGTCCAGCGTATCGATAGCTTGATATCGAATTCCTGCG 900	Qy	1852 GAACATGTGAGCAAAAAGGCCAGCAAAAGGCCAGAAACCGTAANAAGCCCGCTTCTGGC 1911
Db	2805 TGGTATTACAGGTTTCTGTGCGGTAAACCGCAACGGAAGGATGCGCATGTAATGTTCTT 2746	Db	1811 GAACATGTGAGCAAAAAGGCCAGCAAAAGGCCAGAAACCGTAANAAGCCCGCTTCTGGC 1752
Qy	901 AGCCCGGGGATCCACTAGTTTCTAGAGCGCGCGCAACCGCGGTGGAGCTCCAGCTTTGT 960	Qy	1912 GTTTTTTCCATAGGCTCCGCCCTCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAG 1971
Db	2745 TGGGAGAGCGGTTGGCAAGTCAACCGGATCTCTCAATCATATTGGTCAGCAGTTTATACA 2686	Db	1751 GTTTTTTCCATAGGCTCCGCCCTCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAG 1692
Qy	961 TCCCTTTAGTGAGGTTTAAATAGATCCCATGCGCTCAATTTTACGAGACTATCTTTCTAG 1020	Qy	1972 GTGGGAAACCCGACAGGACTATAAAGATACAGGGGTTTTCCCTCTGGAAGCTCCCTCGT 2031
Db	2685 GCTCCGAGAGTCTTTTTCGGTATTCGGATTCGGATAGCCGAAATG--CCCTCGAAATCGGTTG 2628	Db	1691 GTGGGAAACCCGACAGGACTATAAAGATACCAAGGGTTTTCCCTCTGGAAGCTCCCTCGT 1632
Qy		Qy	2032 GCCTCTCTGTTCCGACCTGCGCTTACCGGATACCTGTCGCGCTTTCTCCCTTCGGG 2091
Db		Db	1631 GCCTCTCTGTTCCGACCTGCGCTTACCGGATACCTGTCGCGCTTTCTCCCTTCGGG 1572

2092 AGCGTGGCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCCGTTAGGTGCTTCCG 2151
Db |||||
1571 AAGCGTGGCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCCGTTAGGTGCTTCCG 1512
Qy |||||
2152 CTCGAAGCTGGCTGTGTCAGCAACCCCGTTAGCGCCGACCGCTGCGCTTATCCGG 2211
Db |||||
1511 CTCGAAGCTGGCTGTGTCAGCAACCCCGTTAGCGCCGACCGCTGCGCTTATCCGG 1452
Qy |||||
2212 TAACTATCGTCTTGAGTCCAAACCCGTAAGACACGACTTTATCGCCACTGCGCAGCCAC 2271
Db |||||
1451 TAACTATCGTCTTGAGTCCAAACCCGTAAGACACGACTTTATCGCCACTGCGCAGCCAC 1392
Qy |||||
2272 TGGTAACAGGATTAGCAGACGAGGTATGTAGCGGTGCTACAGATTCTTGAAGTGTG 2331
Db |||||
1391 TGGTAACAGGATTAGCAGACGAGGTATGTAGCGGTGCTACAGATTCTTGAAGTGTG 1332
Qy |||||
2332 GCTTAACCTACGCTACACTAGAAGGACAGTATTTGGTATCTGCGTCTCTCAAGCCAGT 2391
Db |||||
1331 GCTTAACCTACGCTACACTAGAAGGACAGTATTTGGTATCTGCGTCTCTCAAGCCAGT 1272
Qy |||||
2392 TACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACCCCGCTGGTAGCGG 2451
Db |||||
1271 TACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACCCCGCTGGTAGCGG 1212
Qy |||||
2452 TGGTTTTTTTGTGCAACGACGAGATTACGGCGCAGAAAAAAGGATCTCAAGAAGATCC 2511
Db |||||
1211 TGGTTTTTTTGTGCAACGACGAGATTACGGCGCAGAAAAAAGGATCTCAAGAAGATCC 1152
Qy |||||
2512 TTTTGATCTTTTCTACGGGGCTGAGCGCTCAGTGGGAACGAAACTACGTTAAAGGATTTT 2571
Db |||||
1151 TTTTGATCTTTTCTACGGGGCTGAGCGCTCAGTGGGAACGAAACTACGTTAAAGGATTTT 1092
Qy |||||
2572 GGTCTATGAGATTATCAAAAAGGATCTTCACTAGATCTCTTTAAATTTAAATGAAGTTT 2631
Db |||||
1091 GGTCTATGAGATTATCAAAAAGGATCTTCACTAGATCTCTTTAAATTTAAATGAAGTTT 1032
Qy |||||
2632 TAAATCAATCTAAAGTATATATAGTAAACTTTGGTCTGACAGTTACCAATGCTTAAATCAG 2691
Db |||||
1031 TAAATCAATCTAAAGTATATATAGTAAACTTTGGTCTGACAGTTACCAATGCTTAAATCAG 972
Qy |||||
2692 TCAGGACCTATCTCAGCGCATCTGCTATTTTCGTTTCATCATAGTTGCTGACTCCCGT 2751
Db |||||
971 TCAGGACCTATCTCAGCGCATCTGCTATTTTCGTTTCATCATAGTTGCTGACTCCCGT 912
Qy |||||
2752 CGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGCGCCAGTGTGCAATGATACC 2811
Db |||||
911 CGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGCGCCAGTGTGCAATGATACC 852
Qy |||||
2812 GCGAGACCCACGCTCACCGGCTCCAGATTATCAGCAATAAACCAGCCAGCGGAGGGC 2871
Db |||||
851 GCGAGACCCACGCTCACCGGCTCCAGATTATCAGCAATAAACCAGCCAGCGGAGGGC 792
Qy |||||
2872 CGAGCGCAGAGTGGTCTGCAACTTATTCGCGCTCCATCCAGTCTATTAATTTGTCGG 2931
Db |||||
791 CGAGCGCAGAGTGGTCTGCAACTTATTCGCGCTCCATCCAGTCTATTAATTTGTCGG 732
Qy |||||
2932 GGAAGCTAGAGTAAGTAGTTCGCCAGTAAATAGTTTTCGCAACGTTTGGTCCATGCTAC 2991
Db |||||
731 GGAAGCTAGAGTAAGTAGTTCGCCAGTAAATAGTTTTCGCAACGTTTGGTCCATGCTAC 672
Qy |||||
2992 AGGCATCGTGGTGTACCGTCTGCTGGTGGTATGGCTTCATTCAGTCCGGTTCGCAAG 3051
Db |||||
671 AGGCATCGTGGTGTACCGTCTGCTGGTGGTATGGCTTCATTCAGTCCGGTTCGCAAG 612
Qy |||||
3052 ATCAAGGCGAGTTACATGATCCCCATGTTGGTGCAGAAAGGGTTAGCTCTTCGGTCC 3111
Db |||||
611 ATCAAGGCGAGTTACATGATCCCCATGTTGGTGCAGAAAGGGTTAGCTCTTCGGTCC 552
Qy |||||
3112 TCCGATCGTGTGTAGAGTAAGTTGGCGCAGTGTATCACTCATGTTTATGGCAGCACT 3171
Db |||||
551 TCCGATCGTGTGTAGAGTAAGTTGGCGCAGTGTATCACTCATGTTTATGGCAGCACT 492
Qy |||||
3172 GCATAATTTCTTACTGTATGCCATCCGTAAGATGCTTTTCTGTGAATGGTGAGTACTC 3231

Db |||||
491 GCATAATTTCTTACTGTATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTC 432
Qy |||||
3232 AACCAAGTCATTCTCAGAAATAGTGTATGCGCGACCGAGTGTCTCTTGGCCCGCGTCAAT 3291
Db |||||
431 AACCAAGTCATTCTCAGAAATAGTGTATGCGCGACCGAGTGTCTCTTGGCCCGCGTCAAT 372
Qy |||||
3292 ACGGATATAACCGCGCCACATAGCAGAACTTTAAAGTGCTCATCATTTGGAAAAAGTTTC 3351
Db |||||
371 ACGGATATAACCGCGCCACATAGCAGAACTTTAAAGTGCTCATCATTTGGAAAAAGTTTC 312
Qy |||||
3352 TTTGGGGGGAACCTCTCAAGGATCTTACCGCTGTGTAGATCCAGTTCAGTTGTAACCCAC 3411
Db |||||
311 TTTGGGGGGAACCTCTCAAGGATCTTACCGCTGTGTAGATCCAGTTCAGTTGTAACCCAC 252
Qy |||||
3412 TCGTGACCCAACTGATCTTACGATCTTTTACTTTTACAGCGTTTCTGGGTGAGCAAA 3471
Db |||||
251 TCGTGACCCAACTGATCTTACGATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAA 192
Qy |||||
3472 AACAGAAAGGCAAAATGCGCAAAAAGGAATAAGGGCGACACGGAATGTTGAATACT 3531
Db |||||
191 AACAGAAAGGCAAAATGCGCAAAAAGGAATAAGGGCGACACGGAATGTTGAATACT 132
Qy |||||
3532 CATACTCTCTCTTTTCAATATTTTGAAGCATTTATCAGGGTTATGTTCTCATGAGCGG 3591
Db |||||
131 CATACTCTCTCTTTTCAATATTTTGAAGCATTTATCAGGGTTATGTTCTCATGAGCGG 72
Qy |||||
3592 ATACATATTTGAATGTATTTAGAAAAATAAACAATAGGGTTTCCGCGCACATTTCCCGG 3651
Db |||||
71 ATACATATTTGAATGTATTTAGAAAAATAAACAATAGGGTTTCCGCGCACATTTCCCGG 12
Qy |||||
3652 AAAAGTGCCAC 3662
Db |||||
11 AAAAGTGCCAC 1

RESULT 13
CQ768840/c
LOCUS
DEFINITION
Sequence 59 from Patent WO2004007550.
ACCESSION
CQ768840
VERSION
CQ768840.1 GI:45112122
SYNOPSIS
synthetic construct
synthetic construct
other sequences; artificial sequences.
REFERENCE
1
AUTHORS
Unger,C.M., Zehetmeier,C., Torella,C., Niewoehner,J., Ahrens,B. and Beste,G.
TITLE
Mcam inhibitors
JOURNAL
Patent: WO 2004007550-A 59 22-JAN-2004;
Xerion Pharmaceuticals GmbH (DE)
FEATURES
Location/Qualifiers
source
1..3918
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Description of Artificial Sequence: plasmid"

ORIGIN

Query Match 68.8%; Score 2517.8; DB 6; Length 3918;
Best Local Similarity 81.8%; Pred.No. 0;
Matches 3052; Conservative 0; Mismatches 537; Indels 142; Gaps 7;
Qy 3 AAATTGTAAGCGTAAATATTTTGTAAATTCGCGTTAAATTTTGTGTAAATCAGCTCAT 62
Db |||||
3732 AAATTGTAAGCGTAAATATTTTGTAAATTCGCGTTAAATTTTGTGTAAATCAGCTCAT 3673
Qy |||||
63 TTTTAAACCAATAGCGCGAAATCGGCAAAATCCCTTTATAAAATCAAAAGAAATAGACCGAGA 122
Db |||||
3672 TTTTAAACCAATAGCGCGAAATCGGCAAAATCCCTTTATAAAATCAAAAGAAATAGACCGAGA 3613
Qy |||||
123 TAGGTTGAGTGTGTTCCAGTTTGGAAACAGAGTCCACTATTATAAGAACGTTGAGTCTCA 182

Db 3612 TAGGGTTGAGTATTGTTCCAGTTTGGAAACAGAGTCCCACTATTAAAGAAACGTGGACTCGA 3553
QY 183 AGCTCAAAAGGGCGAAACCGTCTATCAGGGCGATGGCCCACTAGCTGAACCATCACCT 242
Db 3552 ACGTCAAAAGGAGAAACCGTCTATCAGGGCGATGGCCCACTAGCTGAACCATCACCA 3493
QY 243 AATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCTTAAGGGAGCC 302
Db 3492 AATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCTTAAGGGAGCC 3433
QY 303 CCCGATTTAGAGCTTGAACGGGAAAGCCGGCGAAGCTGGCGGAGAAAGGAAGAAAG 362
Db 3432 CCCGATTTAGAGCTTGAACGGGAAAGCCGGCGAAGCTGGCGGAGAAAGGAAGAAAG 3373
QY 363 CGAAAGGAGCGGCGCTAGGGCGCTGGCAAGTGTAGCGGTACCGTGGCGGTAAACCA 422
Db 3372 CGAAAGGAGCGGCGCTAGGGCGCTGGCAAGTGTAGCGGTACCGTGGCGGTAAACCA 3313
QY 423 CACCCGCGCGCTTAATGCGCGCTACAGGGCGCT----- 458
Db 3312 CACCGCGCGGCTTAATGCGCGCTACAGGGCGCTACTATGTTGCTTTGACGTATGG 3253
QY 459 -----CCCATTCGCCATT 471
Db 3252 GTGTGAATACCGCACAGATGCGTAAGGAGAAATACCGCATCAGGCGCCATTGCCATT 3193
QY 472 CAGGCTGCGCACTGTTGGGAGGGGATCGGTGGGGCGCTCTCGCTATTATACGCCAGCT 531
Db 3192 CAGGCTGCGCACTGTTGGGAGGGGATCGGTGGGGCGCTCTCGCTATTATACGCCAGCT 3133
QY 532 GCGGAAAGGGGATGTCTGCAAGCGGATTAAGTTGGGTAAAGCCAGGCTTTTCCAGTC 591
Db 3132 GCGGAAAGGGGATGTCTGCAAGCGGATTAAGTTGGGTAAAGCCAGGCTTTTCCAGTC 3073
QY 592 ACGAGCTTTGAAAACGACGGCCAGTGAGCGGCCCTCGTTCACTACGTTTTTGAACCCGT 651
Db 3072 ACGAGCTTTGAAAACGACGGCCAGTGAATTTCTTAAAGACTCTCTTATTAAGCAGTATGT 3013
QY 652 GGAGGACGGCAGACTCCGGGTGCAATGTGTTTACAGCGTGAAGGAGCATGAAGT 711
Db 3012 -----TAGCAACGTCGAAATACATACATAAAGGTGGCAACATATAAAGAA 2964
QY 712 GCTGACACGCTGCAGAACACGACGCTAGATTAACCTAGAAAGATATCATATTGTGAC 771
Db 2963 CGCAAGACACCAACGATAGTTTATTTGTGCAATCAATAGAAATTCATATGGTTT 2904
QY 772 GTACGTTAAAGATATCATGCTAAATTTAGCGATGGGATCTGTAATACGACTCACTAT 831
Db 2903 ACCAGCGCCAAAGACATAAGGGCGACATTCACCCGACTGAGGCAAGAAAGGTAAATATTG 2844
QY 832 AGGCGAATTTGGTACCGGCCCCCTCGAGGTGCGCGGTATCGATAAGCTTGATATCG 891
Db 2843 ACGGAAATTTATCATTAAGGTGAATATFACCCGTACCGACTTGAGCCATTTTGGGAAT 2784
QY 892 AATTCCTGCGAGCCCGGGATCCACTAGTTCTAGAGCGGCCCGCCACCGCGTGGAGCTCC 951
Db 2783 AGAGCCAGCA-----AATCACAGTAGCACCACTTACCATTAGCAAGCCGGAAG 2732
QY 952 AGCTTTTGTCCCTTTAGTAGGGTTAATTTAGATCCCATGGTCAATTTTAGCAGACTA 1011
Db 2731 TCACCAATGAAACCATCGATAGCAGCACCGTAATCAGTAGCGACAGAATCA---AGTTTG 2675
QY 1012 TCTTTCTAGGTTAATCTAGCTGCATCAGGATCATATCGTGGGTCTTTTTCGGCTCA 1071
Db 2674 CTTTAGCTCAGACTGTAGCGCGTTTTCATCGGCAITTTTCGTCATAGCCCTTTATTA 2615
QY 1072 GTCATCGCCAGCTGGCGCTATCTGGGCATCGGGGAGGAAGACCGCTGCTTTTCC 1131
Db 2614 GGTTTGCCATTTTTCATATCAATCAAAATCACCGAAACCGAGCCGCCAAGCGCA 2555
QY 1132 GCGAGTTGAAGCGGCATGGAAGATTTTGGCGAGGATGACTGTGCTGATGACGTG 1191

Db 2554 CCCTCAGAGCCGCCACCTCAGAACCCGCCACCTCAGAGCGGCCACCTCAGAGCGCCA 2495
QY 1192 AGCGAAAAACGACGCTTTTACCATGATGATTCGGGAAGGTGTGGCCATGACGCGCTTTAACG 1251
Db 2494 CAGAAACACAC-----CAGAGCGCGCGCAGCTATGCGGCAAGCGGTTCCAGCGGAT 2440
QY 1252 GTGAACCTGTTGTTTTCAGGCCACCTGGGATAACAGTTTCGTGGGCTTTTTCGGACACAGT 1311
Db 2439 CCGGATACGGCACCGCGCACCTGGCGCGCGCTTGTAGTAGTGCACTTCGCCACCGCC 2380
QY 1312 TCCGATGGTCAGCCGGAAGCGCATCAGAACCCCGAAACAATACCGGCGACAGCCGGAAT 1371
Db 2379 AGAGCACCTTCGGCTGAACCGCCCTCCACCTGTCGA----- 2344
QY 1372 GCGGTGCGGTGTGAGATTAATGACAGCGGTGCGGCTGGGATATTAACGTACAGGAG 1431
Db 2343 -----CTGACCTGGGCCATGGCGGCTGGCGCATAGAAAGAAACAATAAGGNA 2291
QY 1432 ACGGTATCTCTGGCTGGATCGCGAAGAAATGGACATGATACCCCGTGAGTTTACCCGCG 1491
Db 2290 TTGCAATAATAATTTTTCAGTTTGAATACTCCAAAAA-----GGCT 2244
QY 1492 GCGGCGCTTGGGTAATCATGTCATAGCTGTTTCTGTGTGAATTTGTATCCGCTCAC 1551
Db 2243 CCAAGCTTGGCGTAATCATGTCATAGCTGTTTCTGTGAAATTTGTTATCCGCTCAC 2184
QY 1552 AATTCACACACATACGAGCGGGAAGCATAAAGTGTAAAGCTCGGGTGCCTAATAGT 1611
Db 2183 AATTCACACACATACGAGCGGGAAGCATAAAGTGTAAAGCTCGGGTGCCTAATAGT 2124
QY 1612 GAGCTAACTCACTTAATTTGCTGCTCACTGCTCCGCTTTTCCAGTGGGAAAACCTGTC 1671
Db 2123 GAGCTAACTCACTTAATTTGCTGCTCACTGCTCCGCTTTTCCAGTGGGAAAACCTGTC 2064
QY 1672 GTGCCAGCTGCATTAATGAATCGGCAACGCGGGGAGAGCGGTTTTCGCTATTGGGG 1731
Db 2063 GTGCCAGCTGCATTAATGAATCGGCAACGCGGGGAGAGCGGTTTTCGCTATTGGGG 2004
QY 1732 CTCTTCGCTTCTCGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1791
Db 2003 CTCTTCGCTTCTCGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1944
QY 1792 ATCAGCTCACTCAAAAGCGGTAAATACGTTATTCACAGAAATCAGGGGATAACGAGGAAA 1851
Db 1943 ATCAGCTCACTCAAAAGCGGTAAATACGTTATTCACAGAAATCAGGGGATAACGAGGAAA 1884
QY 1852 GAAATGTGAGCAAAAGCGGCAAGAAAGCGGAAACCGTAAGAAAGCGCGCTGCTGCG 1911
Db 1883 GAAATGTGAGCAAAAGCGGCAAGAAAGCGGAAACCGTAAGAAAGCGCGCTGCTGCG 1824
QY 1912 GTTTTTCATAGGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCG 1971
Db 1823 GTTTTTCATAGGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCG 1764
QY 1972 GTGGCAAAACCGACAGACTATAAAGATAACAGCGGTTTCCCTCTGGAAGCTCCCTCGT 2031
Db 1763 GTGGCAAAACCGACAGACTATAAAGATAACAGCGGTTTCCCTCTGGAAGCTCCCTCGT 1704
QY 2032 GCGCTCTCTGTTCCGACCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCG 2091
Db 1703 GCGCTCTCTGTTCCGACCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCG 1644
QY 2092 AAGCTGCGCTTCTCTCATAGCTCAAGCTGATAGGTATCTCAGTTCCGTTGATAGTCTG 2151
Db 1643 AAGCTGCGCTTCTCTCATAGCTCAAGCTGATAGGTATCTCAGTTCCGTTGATAGTCTG 1584
QY 2152 CTCGAAGCTGGGCTGTGTGACGAACCCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCG 2211
Db 1583 CTCGAAGCTGGGCTGTGTGACGAACCCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCG 1524
QY 2212 TAACTATCGTTGAGTCCAAACCGGTAAAGACCACTTATCCCACTTGGCAGCAGCCAC 2271
Db 1523 TAACTATCGTTGAGTCCAAACCGGTAAAGACCACTTATCCCACTTGGCAGCAGCCAC 1464


```
QY 2272 TGATAACAGGATTTAGCAGACGAGGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGGTG 2331
Db 1463 TGATAACAGGATTTAGCAGACGAGGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGGTG 1404
QY 2332 GCCTAACTACGGTACACTAGAGGACAGTATTTGGTATCTCGGCTCTGCTGAAGCCAGT 2391
Db 1403 GCCTAACTACGGTACACTAGAGGACAGTATTTGGTATCTCGGCTCTGCTGAAGCCAGT 1344
QY 2392 TACCTTTGCGAAAAAGAGTTGGTAGTCTTTGATCGCGCAAAACAAACACCGCTGGTAGCGG 2451
Db 1343 TACCTTTGCGAAAAAGAGTTGGTAGTCTTTGATCGCGCAAAACAAACACCGCTGGTAGCGG 1284
QY 2452 TGGTTTTTTTGTGCAAGCAGCAGATTAACGGCAGAAAAAAGAGATCTCAAGAGATCC 2511
Db 1283 TGGTTTTTTTGTGCAAGCAGCAGATTAACGGCAGAAAAAAGAGATCTCAAGAGATCC 1224
QY 2512 TTTGATCTTTTCTACGGGCTCTGACGCTCAGTGGACGAAACTCAGTTAAGGATTTT 2571
Db 1223 TTTGATCTTTTCTACGGGCTCTGACGCTCAGTGGACGAAACTCAGTTAAGGATTTT 1164
QY 2572 GGTATGAGATTAACAAAAAGATCTTACCTAGATCTTTTAAATTAATAATGAAGTTT 2631
Db 1163 GGTATGAGATTAACAAAAAGATCTTACCTAGATCTTTTAAATTAATAATGAAGTTT 1104
QY 2632 TAAATCAATCTAAAGTATATATAGTAAACTTTGGTCTGACAGTTACCAATGCTTAATCAG 2691
Db 1103 TAAATCAATCTAAAGTATATATAGTAAACTTTGGTCTGACAGTTACCAATGCTTAATCAG 1044
QY 2692 TGAGGACCTATCTAGGACATCTGCTATTTGGTTCATCATAGTTGCGCTACTCCCGT 2751
Db 1043 TGAGGACCTATCTAGGACATCTGCTATTTGGTTCATCATAGTTGCGCTACTCCCGT 984
QY 2752 CGTGTAGATAACTACATACGAGGAGGCTTACCATCTGGCCCGCAGTCTGCAATGATACC 2811
Db 983 CGTGTAGATAACTACATACGAGGAGGCTTACCATCTGGCCCGCAGTCTGCAATGATACC 924
QY 2812 GCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCGGAAGGC 2871
Db 923 GCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCGGAAGGC 864
QY 2872 CGAGCGCAGAGTGGTCTGCAACTTTATCGGCTCCATCCAGTCTTAATTTGTCGG 2931
Db 863 CGAGCGCAGAGTGGTCTGCAACTTTATCGGCTCCATCCAGTCTTAATTTGTCGG 804
QY 2932 GGAAGCTAGAGTAAGTAGTTTCGCCAGTAAATAGTTTGGCAAGTGTGGCCATTGCTAC 2991
Db 803 GGAAGCTAGAGTAAGTAGTTTCGCCAGTAAATAGTTTGGCAAGTGTGGCCATTGCTAC 744
QY 2992 AGGCATCGTGGTGTACAGCTCGTGGTGGTATGGCTTCATTACAGCTCCGGTTCCTCAAG 3051
Db 743 AGGCATCGTGGTGTACAGCTCGTGGTGGTATGGCTTCATTACAGCTCCGGTTCCTCAAG 684
QY 3052 ATCAAGCGCAGTATACATGATCCCGCATGTGTGCAAAAAAGCGGTAGCTCTCGGTCC 3111
Db 683 ATCAAGCGCAGTATACATGATCCCGCATGTGTGCAAAAAAGCGGTAGCTCTCGGTCC 624
QY 3112 TCCGATCGTGTGTCAGAAAGTAAGTTGGCCGAGTGTATCACTCATGTTTATGGCAGCAT 3171
Db 623 TCCGATCGTGTGTCAGAAAGTAAGTTGGCCGAGTGTATCACTCATGTTTATGGCAGCAT 564
QY 3172 GCATAATCTCTTACTGTATGTCATCCGATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACT 3231
Db 563 GCATAATCTCTTACTGTATGTCATCCGATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACT 504
QY 3232 AACCAAGTCAATCTGAGAAATAGTGTATGCGGACCGAGTTGCTCTTGGCCCGCGTCAAT 3291
Db 503 AACCAAGTCAATCTGAGAAATAGTGTATGCGGACCGAGTTGCTCTTGGCCCGCGTCAAT 444
QY 3292 ACGGGATAATACCGGCCACATAGCAGAACTTTTAAAGTGCTCATCATTTGGAAAAAGCTTC 3351
Db 443 ACGGGATAATACCGGCCACATAGCAGAACTTTTAAAGTGCTCATCATTTGGAAAAAGCTTC 384
```

```
QY 3352 TTCCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCAC 3411
Db 383 TTCCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCAC 324
QY 3412 TCGTGCAACCAACTGATCTTTCAGCATCTTTTACTTTTACCCAGCGTTTCTGGGTGAGCAAA 3471
Db 323 TCGAGCAACCAACTGATCTTTCAGCATCTTTTACTTTTACCCAGCGTTTCTGGGTGAGCAAA 264
QY 3472 AACAGGAAGCGCAAAATGCGCAAAAGGGAATAAGGGCGACACGGAATAATGTTGAATACT 3531
Db 263 AACAGGAAGCGCAAAATGCGCAAAAGGGAATAAGGGCGACACGGAATAATGTTGAATACT 204
QY 3532 CATACTCTCTCTTTTCAATATATTGAAGCATTTATCAGGGTTATGTTCTCATGAGCGG 3591
Db 203 CATACTCTCTCTTTTCAATATATTGAAGCATTTATCAGGGTTATGTTCTCATGAGCGG 144
QY 3592 ATACATATTTGAATGTTATTAGAAAAATAAACAATAGGGTTCCCGCGACATTTCCCGG 3651
Db 143 ATACATATTTGAATGTTATTAGAAAAATAAACAATAGGGTTCCCGCGACATTTCCCGG 84
QY 3652 AAAAGTGCCAC 3662
Db 83 AAAAGTGCCAC 73

RESULT 14
CQ802124/c
LOCUS 3918 bp DNA linear PAT 05-MAY-2004
DEFINITION Sequence 109 from Patent WO2004007717.
ACCESSION CQ802124
VERSION CQ802124.1 GI:47058654
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Unger, C.M., Beste, G., Jensen, K.H., Zehetmeier, C., Knauer, R.,
Guelbenzu, B.L., Torella, C. and Ilag, L.L.
TITLE Proteomic screen to identify disease-related biological
molecules and inhibitors thereto
JOURNAL Patent: WO 2004007717-A 109 22-JAN-2004;
Xerion Pharmaceuticals AG (DE)
FEATURES
source
1. 3918
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Description of Artificial Sequence: vector pXP10"

ORIGIN

Query Match 68.8%; Score 2517.8; DB 6; Length 3918;
Best Local Similarity 81.8%; Pred. No. 0;
Matches 3052; Conservative 0; Mismatches 537; Indels 142; Gaps 7;

QY 3 AAATTGTAAGCGTTAATAATTTTGTAAAAATTCGCGTTAAATTTTGTAAATCAGCTCAT 62
Db 3732 AAATTGTAAGCGTTAATAATTTTGTAAAAATTCGCGTTAAATTTTGTAAATCAGCTCAT 3673
QY 63 TTTTAAACCAATAGCGCAAAATCGCAAAATCCCTTATAATCAAAAGAAATACACCGAGA 122
Db 3672 TTTTAAACCAATAGCGCAAAATCGCAAAATCCCTTATAATCAAAAGAAATACACCGAGA 3613
QY 123 TAGGGTTGAGTGTGTTCCAGTTTGGAAACAAGAGTCCACTATTATAAGAAACGTGACTCCA 182
Db 3612 TAGGGTTGAGTGTGTTCCAGTTTGGAAACAAGAGTCCACTATTATAAGAAACGTGACTCCA 3553
QY 183 AGCTAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCT 242
Db 3552 AGCTAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCT 3493
QY 243 AATCAAGTTTTTTGGGTCGAGTCCGTAAGACACTTAATCGGAACCTTAAGGGAGGCC 302
Db 3492 AATCAAGTTTTTTGGGTCGAGTCCGTAAGACACTTAATCGGAACCTTAAGGGAGGCC 3433
```

QY 303 CCCGATTTAGAGCTTTGACGGGAAAGCCGCGAAACGTGCGAGAAAGGAAGGAAGAAG 362
DB |||||
3432 CCCGATTTAGAGCTTTGACGGGAAAGCCGCGAAACGTGCGAGAAAGGAAGGAAGAAG 3373
QY 363 CGAAAGGAGCGGCGCTAGGCGCTGGCAAGTGTAGCGGTACAGCTGCGGTAAACCA 422
DB |||||
3372 CGAAAGGAGCGGCGCTAGGCGCTGGCAAGTGTAGCGGTACAGCTGCGGTAAACCA 3313
QY 423 CACCGCGCGCTTAATGCGCGCTACAGGGCGCT----- 458
DB |||||
3312 CACCGCGCGCTTAATGCGCGCTACAGGGCGCTACTATGTTGCTTTGACGTATGCG 3253
QY 459 -----CCATTTCGCCATT 471
DB |||||
3252 GTGTGAATACCGCACAGATGCGTAAGNGAATAACCGCATCAGCGGCCATTTCGCCATT 3193
QY 472 CAGGCTGCGCAACTGTTGGGAAGGCGGATCGGTGCGGCGCTTTCGCTATTACGCCAGCT 531
DB |||||
3192 CAGGCTGCGCAACTGTTGGGAAGGCGGATCGGTGCGGCGCTCTTCGCTATTACGCCAGCT 3133
QY 532 GCGAAAGGGGATGCTGCAAGGCGGATTAAGTTGGGTACGCGCAGGGTTTCCAGTC 591
DB |||||
3132 GCGAAAGGGGATGCTGCAAGGCGGATTAAGTTGGGTACGCGCAGGGTTTCCAGTC 3073
QY 592 ACGAGCTTTGTAAGACGAGCGGCGAGTGAGCGCGCTCGTTTATTACAGTTTGTGAACCCGT 651
DB |||||
3072 ACGAGCTTTGTAAGACGAGCGGCGAGTTGTAATTTTATTAAGACTCTTATTACGCCAGTATGT 3013
QY 652 GGAGGACGGGAGACTCGCGGTGCAAAATGTTGTTTACAGCGTGTAGGAGCAGATGAAGAT 711
DB |||||
3012 -----TAGCAACCGTCGAAATAACATACATAAAGGTGGCAACATATAAAGAAA 2964
QY 712 GCTCGACAGCTGCAGAACCGCAGCTAGATTAAACCTTAGAAGATATCATATTGTGAC 771
DB |||||
2963 CGCAAGACACAACCGGAATAAGTTTATTGTCACAATCAATAGAAATTCATATGGTTT 2904
QY 772 GTACGCTTAAAGATATCATGCTGTAATAATGACGCATGGATCTGTAATACGACTCACTAT 831
DB |||||
2903 ACCAGCGCAAGACATTAAGGCGGACATTCACCGACTGAGGCAAGAAGTAAATATTG 2844
QY 832 AGGGCAATTTGGTTACCGGGCCCCCTCGAGGTGCGAGGTATCGATTAAGCTTGTATGCG 891
DB |||||
2843 ACGGAAATTTATTCATTAAGAGTGAATTTATCACCGTCAACGACTTGAGCCAATTTGGGAAT 2784
QY 892 AATTCTCGACCGCGGGATCCACTAGTTCTAGAGCGGCGGCCACCGGGTGGAGCTCC 951
DB |||||
2783 AGAGCCAGCA-----AAATCACAGTAGCACCAATTACCAATTAGCAAGGCGCGAAACG 2732
QY 952 AGCTTTTGTCCCTTTAGTGAGGGTTAATTAGATCCCATGGTCAATTTTACGCAGACTA 1011
DB |||||
2731 TCACCAATGAACCAATCATAGCAGCACCGTAATCAGTAGCAGACAGATCA---AGTTTG 2675
QY 1012 TCTTTCTAGGTTAATCTAGCTGCAATCAGATCATATCGTGGGTCTTTTTTTCGGCTCA 1071
DB |||||
2674 CTTTATAGGCTCAGACTGTAGCGGCTTTTCATCGGCAITTTTCGGTCATAGCCCTTTATTA 2615
QY 1072 GTCATCGCCCAAGCTGGCGCTATCTGGCATCGGGGAGAGAGCCCGTGCCTTTTCCC 1131
DB |||||
2614 GCGTTTGCATTTTTCATTAATCAAAATACCGGAAACCGGAGCGGCCACCGGAACCGCCA 2555
QY 1132 GCGAGGTTTGAAGCGGCATGGAAGAGTTTGGCGAGGATGACTGTGTCGATTTGACGTTG 1191
DB |||||
2554 CCTCAGAGCGGCCACCTCAGAACCGGCCACCTCAGAGCGGCCACCTCAGAGCGCCA 2495
QY 1192 AGCGAAACCGCAGCTTTTACATGATGATTCGGGAAGGTGTGGCCATGCAAGCTTTTAAG 1251
DB |||||
2494 CGAGAACCAACCA-----CAGAGCGCGCGCCAGCTATGCGGACAGCGGTTCCAGCGGAT 2440
QY 1252 GTGAACTGTTTCAGGCCACCTGGGATACAGTTTCGTGGCGCTTTTCCGGACACAGT 1311
DB |||||
2439 CCGGATACGGCACCGGCGCACCTGCGGCGCGCTTTGATGAGTGACCTTCGCGCACCGCC 2380

QY 1312 TCCGATGCTCAGCCCGAAGCGCATCAGCAACCCGAAACAATATACCGGCGACACCGCGAACT 1371
DB |||||
2379 AGAGCCACCTCGCTCGCTGAACCGCCTCCACCTGTGCA----- 2344
QY 1372 GCCGTGCCGTGTGCAGATTAAATGACAGGGTGCAGCGCTGGGATATTACGTACGGAGG 1431
DB |||||
2343 -----CTGGACCTGGGCCCATGGCGGCTGGCGCGCATAGAAAGGAACAACATAAGGAA 2291
QY 1432 ACGGTATCTCTGGCTGGATGCGCAGAAAATGACATGGATACCCGCTGAGTTACCGCGG 1491
DB |||||
2290 TTGCAATAAATAATTTTTCAGTTGAAAATCTCAAAAAA-----GGCT 2244
QY 1492 GCGCGCTTGGCGTAATCATGCTCATAGCTGTTTCTGTGTGAAAATGTTATTCGCTCAC 1551
DB |||||
2243 CCAAGCTTGGCGTAATCATGCTCATAGCTGTTTCTGTGTGAAAATGTTATTCGCTCAC 2184
QY 1552 AATTCCACAACAATACAGCGCGGAAGCATAAAGTGTAAAGCTTGGGTGCTTAATGAT 1611
DB |||||
2183 AATTCCACAACAATACAGCGCGGAAGCATAAAGTGTAAAGCTTGGGTGCTTAATGAT 2124
QY 1612 GAGCTAACTCACTTAATTTGCGCTTCTGCGCTCACTGCGCTTTCAGTCTGGGAAACCTGTC 1671
DB |||||
2123 GAGCTAACTCACTTAATTTGCGCTCACTGCGCTTTCAGTCTGGGAAACCTGTC 2064
QY 1672 GTGCCAGCTGCAATTAATGAATCGGCCAAACGCGCGGGAGAGCGGTTTTCGCTATTTGGCG 1731
DB |||||
2063 GTGCCAGCTGCAATTAATGAATCGGCCAAACGCGCGGGAGAGCGGTTTTCGCTATTTGGCG 2004
QY 1732 CTCCTCCGCTTCTCGCTCACTGACTCGCTGCGCTCGGTTCGCTGCGGCGAGCGGT 1791
DB |||||
2003 CTCCTCCGCTTCTCGCTCACTGACTCGCTGCGCTCGGTTCGCTGCGGCTGCGGCGAGCGGT 1944
QY 1792 ATCAGCTCACTCAAAAGCGGTAAATACGTTTTCACAGAAATCAGGGGATAACGCAAGGAA 1851
DB |||||
1943 ATCAGCTCACTCAAAAGCGGTAAATACGTTTTCACAGAAATCAGGGGATAACGCAAGGAA 1884
QY 1852 GAACAATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAAACCGTAAAAAGGCGCGCTTCTGCGC 1911
DB |||||
1883 GAACAATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAAACCGTAAAAAGGCGCGCTTCTGCGC 1824
QY 1912 GTTTTTCATAGGCTCCGCGCTTTCAGCAGCATCAAAAATCAGACGCTCAAGTCAAG 1971
DB |||||
1823 GTTTTTCATAGGCTCCGCGCTTTCAGCAGCATCAAAAATCAGACGCTCAAGTCAAG 1764
QY 1972 GTGGGAAACCGCAGGACTATAAAGATACAGGCGCTTTCGCTTTCGAAAGCTCCCTCGT 2031
DB |||||
1763 GTGGGAAACCGCAGGACTATAAAGATACAGGCGCTTTCGCTTTCGAAAGCTCCCTCGT 1704
QY 2032 GCGCTCTCTGTTCCGACCTTGCAGCTTACCGGATACCTGTCGCGCTTTCCTCCCTCGG 2091
DB |||||
1703 GCGCTCTCTGTTCCGACCTTGCAGCTTACCGGATACCTGTCGCGCTTTCCTCCCTCGG 1644
QY 2092 AAGCTGCGGCTTTCATAGCTCAGCTGTAGGTATCTCAGTTGCTGCTAGGTGCTGTCG 2151
DB |||||
1643 AAGCTGCGGCTTTCATAGCTCAGCTGTAGGTATCTCAGTTGCTGCTAGGTGCTGTCG 1584
QY 2152 CTCGAAGCTGCGCTGTGTGACCAACCCCGCTTTCAGGCCAGACGCTGCGCTTATCCG 2211
DB |||||
1583 CTCGAAGCTGCGCTGTGTGACCAACCCCGCTTTCAGGCCAGACGCTGCGCTTATCCG 1524
QY 2212 TAACCTATCGTTGAGTCCAAACCGGTAAGACAGCATTAATGCGCACTGCGGAGCAGCCAC 2271
DB |||||
1523 TAACCTATCGTTGAGTCCAAACCGGTAAGACAGCATTAATGCGCACTGCGGAGCAGCCAC 1464
QY 2272 TGTTAAGAGGATAGCAGAGCGGTATGAGGCGGTCTACAGGTTCTTGAAGTGGTG 2331
DB |||||
1463 TGTTAAGAGGATAGCAGAGCGGTATGAGGCGGTCTACAGGTTCTTGAAGTGGTG 1404
QY 2332 GCCTAACTACGGCTACACTAGAAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAAGCCAGT 2391
DB |||||
1403 GCCTAACTACGGCTACACTAGAAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAAGCCAGT 1344
QY 2392 TACCTTCGGAAGAAAGAGTTGGTAGCTCTTGTATCCGGGCAAAACCAACCGCTGGTAGCGG 2451

Db 1343 TACCTTCGGAAGAGTGTGGTAGCTCTTGATCCGGCAACAAACACCCGCTGGTAGCGG 1284
Qy 2452 TGGTTTTTTTGTGCAAGCAGCAGATTAACGGCGAGAAAAAGGATCTCAAGAGATCC 2511
Db 1283 TGGTTTTTTTGTGCAAGCAGCAGATTAACGGCGAGAAAAAGGATCTCAAGAGATCC 1224
Qy 2512 TTTGATCTTTTCTACGGGCTGTGA CGCTCAGTGGAAACGAAAACTCAGCTTAAGGATTTT 2571
Db 1223 TTTGATCTTTTCTACGGGCTGTGACGCTCAGTGGAAACGAAAACTCAGCTTAAGGATTTT 1164
Qy 2572 GGTGATGAGATTAACAAAGGATCTTACCTAGATCCTTTTAAATTAATAAATGAAGTTT 2631
Db 1163 GGTGATGAGATTAACAAAGGATCTTACCTAGATCCTTTTAAATTAATAAATGAAGTTT 1104
Qy 2632 TAAATCAATCTAAAGTATATAGTAAACTTTGGTCTGACAGTTACCAATGCTTAATCAG 2691
Db 1103 TAAATCAATCTAAAGTATATAGTAAACTTTGGTCTGACAGTTACCAATGCTTAATCAG 1044
Qy 2692 TGAGGCACCTATCTCAGCATCTGTCTATTTGGTTCATCCATAGTTGCTGACTCCCGT 2751
Db 1043 TGAGGCACCTATCTCAGCATCTGTCTATTTGGTTCATCCATAGTTGCTGACTCCCGT 984
Qy 2752 CGTGTAGATAACTACGATACGGAGGCTTACCATCTGCCCGCAGTCTGCAATGATACC 2811
Db 983 CGTGTAGATAACTACGATACGGAGGCTTACCATCTGCCCGCAGTCTGCAATGATACC 924
Qy 2812 GCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCGGAGGCG 2871
Db 923 GCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCGGAGGCG 864
Qy 2872 CGAGCGCAGAGTGGTCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGGCG 2931
Db 863 CGAGCGCAGAGTGGTCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGGCG 804
Qy 2932 GGAAGCTAGAGTAAGTGTGCGCAGTTAATAGTTTGGCAAGCTTTGTCATTTGCTAC 2991
Db 803 GGAAGCTAGAGTAAGTGTGCGCAGTTAATAGTTTGGCAAGCTTTGTCATTTGCTAC 744
Qy 2992 AGGCATCTGTGTGTCAGCTCGTGTGTGTATGGTTCATTCAGCTCCGTTTCCCAAGC 3051
Db 743 AGGCATCTGTGTGTCAGCTCGTGTGTATGGTTCATTCAGCTCCGTTTCCCAAGC 684
Qy 3052 ATCAAGGCGAGTTACATGATCCCCATGTTGTGCAAAAAAGCGGTAGCTCTCGGTCC 3111
Db 683 ATCAAGGCGAGTTACATGATCCCCATGTTGTGCAAAAAAGCGGTAGCTCTCGGTCC 624
Qy 3112 TCCGATCTGTGTCAGAGTAAGTTGGCCGAGTGTATCACTCATGGTTATGGCAGCACT 3171
Db 623 TCCGATCTGTGTCAGAGTAAGTTGGCCGAGTGTATCACTCATGGTTATGGCAGCACT 564
Qy 3172 GCATAATCTCTTACTGTCTATGCTCATCGTAAAGATGCTTTCTGTGACTGGTGAGTACTC 3231
Db 563 GCATAATCTCTTACTGTCTATGCTCATCGTAAAGATGCTTTCTGTGACTGGTGAGTACTC 504
Qy 3232 AACCAAGTCATTTGAGAAATAGTGTATGCGGAGCCAGTGTCTCTTGCCCGCGTCAAT 3291
Db 503 AACCAAGTCATTTGAGAAATAGTGTATGCGGAGCCAGTGTCTCTTGCCCGCGTCAAT 444
Qy 3292 ACGGATTAATACGGGCCACATAGCAGAACTTTAAAGTGTCTCATCATTTGGAAGAGGTTTC 3351
Db 443 ACGGATTAATACGGGCCACATAGCAGAACTTTAAAGTGTCTCATCATTTGGAAGAGGTTTC 384
Qy 3352 TTTCCGGGGGAAATCTCTCAAGGATCTTACCGTGTGTGAGATCCAGTTTCAGTTAAACCCAC 3411
Db 383 TTTCCGGGGGAAATCTCTCAAGGATCTTACCGTGTGTGAGATCCAGTTTCAGTTAAACCCAC 324
Qy 3412 TCGTGACCCAACTGATCTTCAGCATCTTTTACTTTTACCAGGTTTCTGGGTGAGCAAA 3471
Db 323 TCGAGCACCCAACTGATCTTCAGCATCTTTTACTTTTACCAGGTTTCTGGGTGAGCAAA 264
Qy 3472 AACAGAGGCAAAATGCGCGCAAAAAAGGGAATAGGGGCAACGGAAATGTTGAATACT 3531

Db 263 AACAGAGGCAAAATGCGCAAAAGGAATAGGGCGACACACGGAATGTTGAATACT 204
Qy 3532 CATACTCTTCTCTTTTCAATATTAATGAAGCATTTATCAGGGTATTTGTCTCATGAGCGG 3591
Db 203 CATACTCTTCTCTTTTCAATATTAATGAAGCATTTATCAGGGTATTTGTCTCATGAGCGG 144
Qy 3592 ATACATATTTGAATGATTTAGAAAAATAAACAATAGGGTTCGGCGCACATTTCCCGG 3651
Db 143 ATACATATTTGAATGATTTAGAAAAATAAACAATAGGGTTCGGCGCACATTTCCCGG 84
Qy 3652 AAAAGTGCCAC 3662
Db 83 AAAAGTGCCAC 73
RESULT 15
CQ873225/c
LOCUS 3918 bp DNA linear PAT 27-SEP-2004
DEFINITION Sequence 3 from Patent EP1457499.
ACCESSION CQ873225
VERSION CQ873225.1 GI:52746953
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Jay,D.G., Eustace,B.K., Sakurai,T. and Beste,G.
TITLE Inhibitors of extracellular Hsp90
JOURNAL Patent: EP 1457499-A 3 15-SEP-2004;
TUFTS UNIVERSITY SCHOOL OF MEDICINE (US); Xerion Pharmaceuticals AG
(DE)
FEATURES
source
1. .3918
/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"
ORIGIN
Query Match 68.8%; Score 2517.8; DB 6; Length 3918;
Best Local Similarity 81.8%; Pred. No. 0;
Matches 3052; Conservative 0; Mismatches 537; Indels 142; Gaps 7;
Qy 3 AAATTGTAAAGCTTAATATTTTGTAAATTCGCGTAAATTTTGTAAATCAGTCAAT 62
Db 3732 AAATTGTAAAGCTTAATATTTTGTAAATTCGCGTAAATTTTGTAAATCAGTCAAT 3673
Qy 63 TTTTAAACCAATAGCGCGAAATCGGCATAATCCCTTATAAATCAAGAATAGACCGAGA 122
Db 3672 TTTTAAACCAATAGCGCGAAATCGGCATAATCCCTTATAAATCAAGAATAGACCGAGA 3613
Qy 123 TAGGGTGTAGTGTGTCTCCAGTTTGGAAACAGAGTCCACTATTAAAGAACGTGGACTCCA 182
Db 3612 TAGGGTGTAGTGTGTCTCCAGTTTGGAAACAGAGTCCACTATTAAAGAACGTGGACTCCA 3553
Qy 183 ACGTCAAAAGGGCGAAAAACCGTCTATCAGGGCGAATGGCCCACTAGCTGAACCAATCACT 242
Db 3552 ACGTCAAAAGGGCGAAAAACCGTCTATCAGGGCGAATGGCCCACTAGCTGAACCAATCACT 3493
Qy 243 AATCAAGTTTTTGGGGTTCGAGTCCGTAAAGCACTAAATCGGAACCTTAAGGGAGCC 302
Db 3492 AATCAAGTTTTTGGGGTTCGAGTCCGTAAAGCACTAAATCGGAACCTTAAGGGAGCC 3433
Qy 303 CCCGATTTTAGACTTTGACCGGGAAAGCCGGCGAAACGTGGCGAAGAGGAAGGAAGAAAG 362
Db 3432 CCCGATTTTAGACTTTGACCGGGAAAGCCGGCGAAACGTGGCGAAGAGGAAGGAAGAAAG 3373
Qy 363 CGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACTGCGTGCCTAACCAACA 422
Db 3372 CGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACTGCGTGCCTAACCAACA 3313
Qy 423 CACCCGCGCGCTTAATGCGCGCTTACAGGGCGCGT----- 458

Db 3312 CACCCGCGGGCTTAATGCGCGCTACAGGCGCGTACTATGCTTGCTTTCAGCTATGCG 3253
Qy 459 -----CCCATTCGCCATT 471
Db 3252 GTGTGAATATACCGCACAGATGCGTAAGGAGAAAAATACCGCATCAGGCGCATTCGCCATT 3193
Qy 472 CAGGCTGGCGCAACTGTGTGGGAAGGCGATCGGTGGGGGCTCTTCGCTATTACGCCAGCT 531
Db 3192 CAGGCTGGCGCAACTGTGTGGGAAGGCGATCGGTGGGGGCTCTTCGCTATTACGCCAGCT 3133
Qy 532 GCGAAGAGGGGATGTGTCGAAGCGGATTAAGTTGGGTAAAGCCAGGTTTTCAGATC 591
Db 3132 GCGAAGAGGGGATGTGTCGAAGCGGATTAAGTTGGGTAAAGCCAGGTTTTCAGATC 3073
Qy 592 ACGAGCTGTAAAGACGAGCGGCAGTGAGCGCGCTCGCTTCATTCAGCTTTTGAACCCGT 651
Db 3072 ACGAGCTGTAAAGACGAGCGGCAGTGGAATCTTATTAGAGCTCCTTATTACCGCATATGT 3013
Qy 652 GGAGGACGGGCAAGTCTCGCGGTGCAAAATGTGTTTACAGCGTGAATGGAGCAGATGAAGAT 711
Db 3012 -----TAGCAAAAGCTCGAATAATACATACATAAAGGTGGCAACATATAAAGAAA 2964
Qy 712 GCTCGACAGCTGCAGAACACCGCAGCTAGATTAACTCTAGAAAGATTAATCATATTGTGAC 771
Db 2963 CGCAAGAGCACCAACGGAATAAGTTTATTGTCACAACTCAATAGAAAAATTCATATGGTTT 2904
Qy 772 GTACGTTAAAGATATCATGCTAAATTTGACGATGGGATCTGTAATACGACTCACTAT 831
Db 2903 ACCAGCGCAAGACATTAAGGGCGACATTCACCGCATGAGGCAAAAGGTAATAATG 2844
Qy 832 AGGCGCAATTTGGGTACCGGGGCCCGCTCGAGGTGCGAGGTATCGAATAGCTTTGATATCG 891
Db 2843 ACGGAATATTTCATTAAGAGTGAATATACCGCTACCGCATTTGGGAATT 2784
Qy 892 AATTCCTGAGCCCGGGGATCCATAGTTCTAGAGCGGCGGCCACCGCGGTGAGCTCC 951
Db 2783 AGAGCGCAGCA-----AAATCACCAGTAGCACCAATTACCAATTAGCAAGGCGGAAAAG 2732
Qy 952 AGCTTTTGTTCCTTTAGTGAGGGTTAATTAGATCCCATGCGTCAATTTTACGACAGCTA 1011
Db 2731 TCACCAATGAACCAATCGATAGCAGCACCGTAAATCAGTAGCAGACAGATCA---AGTTTG 2675
Qy 1012 TCTTTCTAGGGTTAATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGGTCA 1071
Db 2674 CTTTATAGGTGAGCTGATAGCGGTTTTCATCGGCAATTTTCGCTCATAGCCCTTAITTA 2615
Qy 1072 GTCATCGCCAGCTGCGGTCTATCTGGGATCGGGAGGAGAACCGCGTCTTTTCCC 1131
Db 2614 GCGTTTGCCATTTTTCATAATCAAAATCACCGGAACCGGAGCCGCCACCGGAACCGCCA 2555
Qy 1132 GCGAGTTGAAGCGGCATGGAAGAGTTTGGCGAGGATGACTGCTGCTGATTTGACGTTG 1191
Db 2554 CCTCAGAGCCGCCACCTCAGAACCGGCACCCCTCAGAGCCGCCACCTCAGAGCGCCA 2495
Qy 1192 AGCGAAAAAGCGAGTTTACCATGATGATTGCGGAAGGTGTGGCCATGCAAGCTTTAAAG 1251
Db 2494 CCAGAACCCACAC-----CAGAGCCGCGCAGCCTATGCGGCACCGGTTCCAGCGGAT 2440
Qy 1252 GTGAATGTTCTGTCAGGCCACCTGCGGATACAGTTTCTGCGGCTTTTTCGGGACACAGT 1311
Db 2439 CCGGATACCGCACCCGCGCACCTGCGGCGCGGTTTGTATGATGCACTTTCCGCCACCGCC 2380
Qy 1312 TCCGATGTCAGCGGAGCGCATCAGCAACCGCAACATACCGGCGAGACCGGAACT 1371
Db 2379 AGAGCCACCTCCGCTGAAACCGCCCTCCACTGTGCA----- 2344
Qy 1372 GCGGTGCGGCTGTGAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTACGAGG 1431
Db 2343 -.-.-.-.-CTGGACCTGGGCCATAGCGCGGTGCGGCGCATAGAAAGAAACAACTAAAGGAA 2291
Qy 1432 ACGGATATCTCTGCTGATGCGGAGAAATGACATGGAATACCGGTGAGTTACCGGG 1491
Db 2290 TTGCGAATTAATTAATTTTTCAGTTTGAAATCTCCAAAAA-----GGCT 2244

Qy 1492 GCGCGCTTGGCGTAATCATGCTCATAGCTGTTTCTCTGTGAAATTTGTTATCCGCTCAC 1551
Db 2243 CCAAGCTTGGCGTAATCATGCTCATAGCTGTTTCTCTGTGAAATTTGTTATCCGCTCAC 2184
Qy 1552 AATTCACACAAATACATACAGCGCGGAAGCATAAAGTGTAAAGCTGGGGTGCCTAATAGT 1611
Db 2183 AATTCACACAAATACATACAGCGCGGAAGCATAAAGTGTAAAGCTGGGGTGCCTAATAGT 2124
Qy 1612 GAGCTAATCTCATTAATTTGCTGCGCTCATGCGCCGCTTTCAGTCTGGGAAACCTGTC 1671
Db 2123 GAGCTAATCTCATTAATTTGCTGCGCTCATGCGCCGCTTTCAGTCTGGGAAACCTGTC 2064
Qy 1672 GTGCCAGCTGCAATTAATGAATCGGCCAAACCGCGCGGAGAGCGGTGTTGCGTATTGGGCG 1731
Db 2063 GTGCCAGCTGCAATTAATGAATCGGCCAAACCGCGCGGAGAGCGGTGTTGCGTATTGGGCG 2004
Qy 1732 CTCTTTCGCTTCTCGCTCATGCTGCGTCTGCGTCTGCTGCTGCTGCGGCGAGCGGT 1791
Db 2003 CTCTTTCGCTTCTCGCTCATGCTGCGTCTGCGTCTGCTGCTGCGGCGAGCGGT 1944
Qy 1792 ATCAGCTCACTCAAGGCGGTAAATACGGTTATCCACAGAACTCAGGCGATAAACGAGGAAA 1851
Db 1943 ATCAGCTCACTCAAGGCGGTAAATACGGTTATCCACAGAACTCAGGCGATAAACGAGGAAA 1884
Qy 1852 GAAATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAAACCGTAAAAAGGCCGCTTGTGTCG 1911
Db 1883 GAAATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAAACCGTAAAAAGGCCGCTTGTGTCG 1824
Qy 1912 GTTTTTCATAGGCTCCCGCCCTCTGACAGCATCAAAAAATCGACGCTCAAGTCAGAG 1971
Db 1823 GTTTTTCATAGGCTCCCGCCCTCTGACAGCATCAAAAAATCGACGCTCAAGTCAGAG 1764
Qy 1972 GTGGCGAAACCGCAGACGATATAAGATACAGGCGTTTCCCGCTGGAAGCTCCCTCGT 2031
Db 1763 GTGGCGAAACCGCAGACGATATAAGATACAGGCGTTTCCCGCTGGAAGCTCCCTCGT 1704
Qy 2032 GCGCTCTCTGTTCCGACCCCTGCGCTTACCGGATACCTGTCCGCTTTTCTCCCTTCGGG 2091
Db 1703 GCGCTCTCTGTTCCGACCCCTGCGCTTACCGGATACCTGTCCGCTTTTCTCCCTTCGGG 1644
Qy 2092 AAGCTGGCGCTTCTCATAGCTCAGCTGATAGTATCTCAGTTGCTGCTGATAGTCTGCTG 2151
Db 1643 AAGCTGGCGCTTCTCATAGCTCAGCTGATAGTATCTCAGTTGCTGCTGATAGTCTGCTG 1584
Qy 2152 CTCCAAGCTGGGCTGTGTGCAGAACCCCGCTTACAGCCCGCTGCGCTTATCCGG 2211
Db 1584 CTCCAAGCTGGGCTGTGTGCAGAACCCCGCTTACAGCCCGCTGCGCTTATCCGG 1524
Qy 2212 TAACTATCTGTTGAGTCCAAACCCCGTAAAGACACGACTTATCCCACTGCGCAGAGCCAC 2271
Db 1523 TAACTATCTGTTGAGTCCAAACCCCGTAAAGACACGACTTATCCCACTGCGCAGAGCCAC 1464
Qy 2272 TGGTAACAGGATTACAGAGGAGTATGAGGCGGTGCTACAGAGTCTTGAAGTGGT 2331
Db 1463 TGGTAACAGGATTACAGAGGAGTATGAGGCGGTGCTACAGAGTCTTGAAGTGGT 1404
Qy 2332 GCTAATCTAGGCTACACTAGAGGACAGTATTTGCTATCTGCGCTCTGCTGAAGCCAGT 2391
Db 1403 GCTAATCTAGGCTACACTAGAGGACAGTATTTGCTATCTGCGCTCTGCTGAAGCCAGT 1344
Qy 2392 TACCTTCGGAAAAAGAGTTGGTAGCTCTTGTATCCGGCAAAACCAACCCGCTGTTAGCGG 2451
Db 1343 TACCTTCGGAAAAAGAGTTGGTAGCTCTTGTATCCGGCAAAACCAACCCGCTGTTAGCGG 1284
Qy 2452 TGGTTTTTTTGTTCGAAGAGCAGATTAACGCGCAGAAAAAAGGATCTCAAGAGATCC 2511
Db 1283 TGGTTTTTTTGTTCGAAGAGCAGATTAACGCGCAGAAAAAAGGATCTCAAGAGATCC 1224
Qy 2512 TTTGATCTTTTCTACGGGCTGAGCGCTCAGTGGAAACGAAAACTCACTGTTAAGGGATTTT 2571
Db 1223 TTTGATCTTTTCTACGGGCTGAGCGCTCAGTGGAAACGAAAACTCAGCTTAAGGGATTTT 1164

```
QY 2572 GGTCTAGAGTATCAAAAGGATCTTACCTAGATCCTTTTAAATTAATAATCAAGTTT 2631
Db 1163 GGTCTAGAGTATCAAAAGGATCTTACCTAGATCCTTTTAAATTAATAATCAAGTTT 1104
QY 2632 TAAATCAATCAAAAGTATATAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAG 2691
Db 1103 TAAATCAATCAAAAGTATATAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAG 1044
QY 2692 TGAGGCACTTATCTCAGCGATCTGTCTATTTCTTTTATCCATAGTTGCCCTGACTCCCCGT 2751
Db 1043 TGAGGCACTTATCTCAGCGATCTGTCTATTTCTTTTATCCATAGTTGCCCTGACTCCCCGT 984
QY 2752 CGTGTAGATAACTACGAGGAGGCTTACCATCTGGCCCCAGTGTGCAATGATACC 2811
Db 983 CGTGTAGATAACTACGAGGAGGCTTACCATCTGGCCCCAGTGTGCAATGATACC 924
QY 2812 GCAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGC 2871
Db 923 GCAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGC 864
QY 2872 CGAGCCGAGAAGTGGTCTTGCAACTTTATCCGGCTCCATCCAGTCTATTAAATTTGCCG 2931
Db 863 CGAGCCGAGAAGTGGTCTTGCAACTTTATCCGGCTCCATCCAGTCTATTAAATTTGCCG 804
QY 2932 GGAAGCTAGAGTAAGTAGTTCCGCCAGTTAATAGTTTGGCCAAAGTTTGGCCATTGCTAC 2991
Db 803 GGAAGCTAGAGTAAGTAGTTCCGCCAGTTAATAGTTTGGCCAAAGTTTGGCCATTGCTAC 744
QY 2992 AGGCATCGTGGTGTACGCTCGTGTGGTATGGCTTCAATCAGCTCCGGTTCCCAACG 3051
Db 743 AGGCATCGTGGTGTACGCTCGTGTGGTATGGCTTCAATCAGCTCCGGTTCCCAACG 684
QY 3052 ATCAAGCGAGTTACATGATCCCCATGTTGTGCAAAAGCGGTTAGCTCTTCCGGTCC 3111
Db 683 ATCAAGCGAGTTACATGATCCCCATGTTGTGCAAAAGCGGTTAGCTCTTCCGGTCC 624
QY 3112 TCCGATCGTGTGTCAGAAAGTGGTGGCCGAGTGTATCACTCATGTTATGSCAGCACT 3171
Db 623 TCCGATCGTGTGTCAGAAAGTGGTGGCCGAGTGTATCACTCATGTTATGSCAGCACT 564
QY 3172 GCATAATCTCTTACTGTCTATGCTATCCGATCCGATGCTTCTTCTGCTGCTGCTGCTGCT 3231
Db 563 GCATAATCTCTTACTGTCTATGCTATCCGATCCGATGCTTCTTCTGCTGCTGCTGCTGCT 504
QY 3232 AACCAAGTCAATCTGAGAAATAGTGTATGCGGCGACCGAGTGTCTTGTGCCCGCGTCAAT 3291
Db 503 AACCAAGTCAATCTGAGAAATAGTGTATGCGGCGACCGAGTGTCTTGTGCCCGCGTCAAT 444
QY 3292 ACGGGATATACCGCGCCACATAGCAGAACTTTAAAGTGTCTCATCATTTGGAAACGTTTC 3351
Db 443 ACGGGATATACCGCGCCACATAGCAGAACTTTAAAGTGTCTCATCATTTGGAAACGTTTC 384
QY 3352 TTCGGGGCGAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAAACCCAC 3411
Db 383 TTCGGGGCGAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAAACCCAC 324
QY 3412 TCGTGACCCCACTGATCTTTCAGCATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAA 3471
Db 323 TCGTGACCCCACTGATCTTTCAGCATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAA 264
QY 3472 AACAGGAAGCAAAATCGCGCAAAATAAGGAATTAAGGGCGACACGGAAATGTTGAATACT 3531
Db 263 AACAGGAAGCAAAATCGCGCAAAATAAGGAATTAAGGGCGACACGGAAATGTTGAATACT 204
QY 3532 CATACCTCTCTTTTCAATTTATTTGAGCATTTTATCAGGTTTATGCTCATGAGCGG 3591
Db 203 CATACCTCTCTTTTCAATTTATTTGAGCATTTTATCAGGTTTATGCTCATGAGCGG 144
QY 3592 ATACATATTTGAATGTATTTAGAAAAATAAACAATAGGGGTTCCGCGCACATTTCCCGG 3651
Db 143 ATACATATTTGAATGTATTTAGAAAAATAAACAATAGGGGTTCCGCGCACATTTCCCGG 84
QY 3652 AAAAGTGCCAC 3662
```

Db 83 AAAAGTGCCAC 73

Search completed: June 5, 2005, 19:22:50
Job time : 15400.2 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 5, 2005, 13:34:58 ; Search time 1772.75 Seconds
(without alignments)
12228.534 Million cell updates/sec

Title: US-10-826-523-41

Perfect score: 3662

Sequence: 1 ccaattgtaagcgtaata.....atttccccgaaagtgccac 3662

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3648.4	99.6	3661	10 ACA55354	AcA55354 Transform
C	2665.4	72.8	3637	11 ADM68463	Adm68463 Variant g
C	2665.4	72.8	3637	12 ADP26616	Adp26616 Green flu
C	2665.4	72.8	3637	13 ADQ88438	Adq88438 pBSC3BFP
C	2665.4	72.8	3637	13 ADS17192	Ada17192 pBSC3BFP
C	2664.8	72.8	3637	8 ABX14571	Abx14571 Plasmid p
C	2664.8	72.8	3637	11 ADM68448	Adm68448 Cycle 3 g
C	2664.8	72.8	3637	12 ADP26601	Adp26601 Green flu
C	2664.8	72.8	3637	13 ADQ88423	Adq88423 pBSC3GFP
C	2664.8	72.8	3637	13 ADS17177	Ada17177 pBSC3BFP
C	2656.4	72.5	3637	8 ABX14570	Abx14570 Plasmid p
C	2656.4	72.5	3637	11 ADM68462	Adm68462 Wild type
C	2656.4	72.5	3637	12 ADP26615	Adp26615 Green flu
C	2656.4	72.5	3637	13 ADQ88437	Adq88437 pBSC3GFP
C	2656.4	72.5	3637	13 ADS17191	Ada17191 pBSC3GFP
C	2647.8	72.3	3637	10 ADI26329	Adi26329 Novel end
C	2635.8	72.0	3681	2 AAQ13578	Aaq13578 Plasmid p
C	2622.7	71.6	5314	2 AAQ13576	Aaq13576 Plasmid p
C	2620.6	71.6	3699	2 AAQ14340	Aaq14340 Plasmid p
C	2609.4	71.3	4001	6 AAD27066	Aad27066 Plasmid p

C	21	2517.8	68.8	3918	13	ADR70323	Adr70323 Single ch
C	22	2517.8	68.8	3918	13	ADS17527	Adsl7527 Nucleotid
	23	2484.6	67.8	3774	6	AAAD27062	Aad27062 Plasmid p
	24	2474.2	67.6	5277	3	AAA88110	Aaa88110 Plasmid p
	25	2464.6	67.3	3908	6	ABT08165	Abt08165 Recombina
	26	2451.8	67.0	3928	8	ABT14478	Abt14478 HCV envel
	27	2451.8	67.0	3928	8	AAAD50628	Aad50628 pBKS-E26H
	28	2451.8	67.0	3928	10	ABQ84236	Abq84236 Vector pB
	29	2448.8	66.9	3956	2	AAV64258	Avv64258 Plasmid p
	30	2446.8	66.8	3715	6	AAAD27064	Aad27064 Plasmid T
	31	2446.8	66.8	3927	6	ABT08166	Abt08166 Recombina
	32	2314.8	63.2	4088	2	AAV64255	Avv64255 Plasmid p
	33	2305.2	62.9	4102	2	AAV64257	Avv64257 Plasmid p
	34	2279.2	62.2	4119	6	AAAD40772	Aad40772 Plasmid p
	35	2278	62.2	3143	13	ADQ88298	Adq88298 High copy
	36	2271.2	62.0	4059	12	ADN97131	Adn97131 Mammalian
	37	2230.6	60.9	5250	10	AAAD49957	Aad49957 TAR (Tran
	38	2224.6	60.7	2958	2	AAZ22250	Aaz22250 Nucleotid
	39	2216.2	60.5	4205	4	AAAD09088	Aad09088 PEAG658 p
	40	2216.2	60.5	4205	6	AAK99713	Aak99713 DNA of pl
C	41	2215.4	60.4	4172	2	AAAX19901	Aax19901 Plasmid p
	42	2213	60.4	4205	4	AAAD09087	Aad09087 PEAG657 p
	43	2213	60.4	4205	6	AAK99700	Aak99700 DNA of pl
	44	2202.6	60.1	4454	4	AAAD06386	Aad06386 Vector pG
	45	2172.8	59.3	7102	2	AAAX1836	Aax1836 Nucleotid

ALIGNMENTS

RESULT 1

ACA55354
ID ACA55354 standard; DNA; 3661 BP.

XX ACA55354;

XX 06-JUN-2003 (first entry)

XX Transformation vector piggyBAC related plasmid pXL-Bac.
XX PiggyBac; transposon; eukaryotic transformation vector; ds;
XX transformed cell; transformed embryo; transgenic; plasmid; cyclic;
XX circular.

XX Synthetic.

XX US2002173634-A1.

XX 21-NOV-2002.

XX 30-OCT-2001; 2001US-00001189.

XX 31-OCT-2000; 2000US-0244677P.

XX 01-NOV-2000; 2000US-0244984P.

XX (FRAS/) FRASER M J.

XX (LIXX/) LI X.

XX (BEAM/) BEAM T.

XX (HUA/) HUA-VAN A.

XX Fraser MJ, Li X, Beam T, Hua-Van A;

XX WPI; 2003-352597/33.

XX New DNA molecule in the transposon piggyBac, useful for transferring genes into host cells or embryos for transforming the cells of embryos that can be used in making transgenic organisms.
XX Example 3; Fig 3(C2); 151pp; English.

XX The invention describes a DNA molecule comprising at least 163 consecutive nucleotide base pairs of the 3' terminal region beginning at the 3' terminal base pair, and at least 125 consecutive nucleotide base

CC pairs of the 5' terminal region beginning at the 5'terminal base pair, of
CC the piggybac molecule. The region extends from the restriction site SacI
CC to the end of the piggybac molecule. The DNA molecule in the transposon
CC piggybac is useful for transferring genes into host cells or embryos for
CC transforming the cells of embryos. The transformed cells or embryos are
CC useful for developing or making transgenic organisms. This sequence
CC represents plasmid used in the creation of minimal sequence eukaryotic
CC transformation vector piggyBac

```
Query Match      99.6%; Score 3648.4; DB 10; Length 3661;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3660; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
```

Qy	1	CTAAATGTTAAGCGTTAATAATTTTGTGTTAAAAATTCGGTTTAAATTTTGTGTTAAATCAGCTC	60
Db	1	CTAAATGTTAAGCGTTAATAATTTTGTGTTAAAAATTCGGTTTAAATTTTGTGTTAAATCAGCTC	60
Qy	61	ATTTTTTAAACCAATAGGCCGAAATCGGCAAAATCCCTTATAATCAAAAGAATAGACCGA	120
Db	61	ATTTTTTAAACCAATAGGCCG- AATCGGCAAAATCCCTTATAATCAAAAGAATAGACCGA	119
Qy	121	GATAGGGTTGAGTGTGTTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGAAACGTGCACTC	180
Db	120	GATAGGGTTGAGTGTGTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGAAACGTGCACTC	179
Qy	181	CAACGTCAAAAGGGGGAAGAAAACGCTCTATCAGGGCGATGGCCCACTACGTGAACCATCAC	240
Db	180	CAACGTCAAAAGGGGGAAGAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCAC	239
Qy	241	CTAATCAAGTTTTTTTGGGGTTCGAGGTGCGTAAAGCACATAATCGAAACCCCTAAAGGGAG	300
Db	240	CTAATCAAGTTTTTTTGGGGTTCGAGGTGCGTAAAGCATTAATTCGAAACCCCTAAAGGGAG	299
Qy	301	CCCCCGATTTAGAGCTTGACCGGGGAAAGCCGCGCAACGTGGCGAGAAAGAAAGGAAAGAA	360
Db	300	CCCCCGATTTAGAGCTTGACCGGGGAAAGCCGCGCAACGTGGCGAGAAAGGAAAGGAAAGAA	359
Qy	361	AGCGAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGGGTCACTGCTGGCGGTAAACAC	420
Db	360	AGCGAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTACGCTGGCGGTAAACAC	419
Qy	421	CACACCCGCGCGCTTAATGCGCGCTACAGGGCGCTCCCAATTCGCCAATTCAGAGCTGCG	480
Db	420	CACACCCGCGCGCTTAATGCGCGCTAAGGGCGCTGCCAATTCGCCAATTCAGAGCTGCG	479
Qy	481	CAACTGTTGGAAAGGGCGATCGGTGCGGCGCTCTTCGTATTACGCCAGCTGGCGAAAGG	540
Db	480	CAACTGTTGGAAAGGGCGATCGGTGCGGCGCTCTTCGTATTACGCCAGCTGGCGAAAGG	539
Qy	541	GGGATGTGCTCAAGGCGATTAAGTTGGGTAAACGCGCAGGGTTTTCCAGTACAGAGTTG	600
Db	540	GGGATGTGCTCAAGGCGATTAAGTTGGGTAAACGCGCAGGGTTTTCCAGTACAGAGTTG	599
Qy	601	TAAACGACGCGCAGTAGCGCGCTCGTTCAATTCACAGTTTTTGAACCCGTCGAGAGCG	660
Db	600	TAAACGACGCGCAGTAGCGCGCTCGTTCAATTCACAGTTTTTGAACCCGTCGAGAGAGCG	659
Qy	661	GCAGACTCGCGGTCAAAATGTGTTTTACAGCGTGA TGGAGCAGATCAAGATGCTCGACAC	720
Db	660	GCAGACTCGCGGTCAAAATGTGTTTTACAGCGTGATGGAGCAGATCAAGATGCTCGACAC	719
Qy	721	GCTCGAGAACAGCGAGCTAGATTAAACCTAGAAAGATAATCATATTGTGACGTACGTTAA	780
Db	720	GCTCGAGAACAGCGAGCTAGATTAAACCTAGAAAGATAATCATATTGTGACGTACGTTAA	779
Qy	781	AGATAATCATCGTAAATTCACGCATCGGATCTGTAATACGACTCACTATATAGGCGCAAT	840
Db	780	AGATAATCATCGTAAATTCACGCATCGGATCTGTAATACGACTCACTATATAGGCGCAAT	839
Qy	841	TGGGTACCGGGCCCCCCTTCGAGGTTCGAGGTATCGATAAGCTTCGATATCGAATTCCTGCG	900

```
QY 1981 CCGGACAGGACTATAAGATACACAGCGGTTTCCCTTGGAGCTCCCTCGTGGCTCTCC 2040
Db 1980 CCGGACAGGACTATAAGATACACAGCGGTTTCCCTTGGAGCTCCCTCGTGGCTCTCC 2039
QY 2041 TGTTCGAGCCCTCGCGCTTACCGGATACCTGTCCGCTTCTCCCTTCGGGAGCGTGGC 2100
Db 2040 TGTTCGAGCCCTCGCGCTTACCGGATACCTGTCCGCTTCTCCCTTCGGGAGCGTGGC 2099
QY 2101 GCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCCGGTGTAGGTGGTTCGCTCCAAGCT 2160
Db 2100 GCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCCGGTGTAGGTGGTTCGCTCCAAGCT 2159
QY 2161 GGGCTGTGTGCAAGAACCCCGCTTACGCCGATCCGCTGCGCTTATCCGGTAACTATCG 2220
Db 2160 GGGCTGTGTGCAAGAACCCCGCTTACGCCGATCCGCTGCGCTTATCCGGTAACTATCG 2219
QY 2221 TCTTGAGTCCAAACCCGGTAAGACACGACTTATCGCCACTGCGCAGCGCACTGGTAACAG 2280
Db 2220 TCTTGAGTCCAAACCCGGTAAGACACGACTTATCGCCACTGCGCAGCGCACTGGTAACAG 2279
QY 2281 GATTACGAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTAACTA 2340
Db 2280 GATTACGAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTAACTA 2339
QY 2341 CGGCTACACTAGAAAGACAGTATTTGGTATCTCGCTCTGCTGAAAGCCAGTTACCTCGG 2400
Db 2340 CGGCTACACTAGAAAGACAGTATTTGGTATCTCGCTCTGCTGAAAGCCAGTTACCTCGG 2399
QY 2401 AAAAAAGTGTGTAGCTCTTGATCCGGCAAAACAAACCCGCTGGTAGCGGTGTTTTT 2460
Db 2400 AAAAAAGTGTGTAGCTCTTGATCCGGCAAAACAAACCCGCTGGTAGCGGTGTTTTT 2459
QY 2461 TGTTCGACAGCAGATACGCGCAGAAAAAAGGATCTCAAGAGATCTCTTGATCTT 2520
Db 2460 TGTTCGACAGCAGATACGCGCAGAAAAAAGGATCTCAAGAGATCTCTTGATCTT 2519
QY 2521 TTCTACGGGCTGTACGCTCAGTGGAAACGAAACTCACGTTAAGGATTTGGTCAATGAG 2580
Db 2520 TTCTACGGGCTGTACGCTCAGTGGAAACGAAACTCACGTTAAGGATTTGGTCAATGAG 2579
QY 2581 ATTATCAAAAAGGATCTTACCTAGATCTCTTAAATTAAGATGAATTTAAATCAAT 2640
Db 2580 ATTATCAAAAAGGATCTTACCTAGATCTCTTAAATTAAGATGAATTTAAATCAAT 2639
QY 2641 CTAAAGTATATAGTAAACTTGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACC 2700
Db 2640 CTAAAGTATATAGTAAACTTGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACC 2699
QY 2701 TATCTCAGCGATCTGTCTATTTTCGTTTATCCATATAGTTGCTGACTCCCCGTGCTGAT 2760
Db 2700 TATCTCAGCGATCTGTCTATTTTCGTTTATCCATATAGTTGCTGACTCCCCGTGCTGAT 2759
QY 2761 AACTAGATACGGAGGGCTTACCATCTGGCCCCAGTGTGCAATATACCGGAGACC 2820
Db 2760 AACTAGATACGGAGGGCTTACCATCTGGCCCCAGTGTGCAATATACCGGAGACC 2819
QY 2821 ACCTCACCGGCTCCAGATTATCAGCAATAAACACGAGCGAGCGGAGCGGCGAG 2880
Db 2820 ACCTCACCGGCTCCAGATTATCAGCAATAAACACGAGCGAGCGGAGCGGCGAG 2879
QY 2881 AAGTGTCTCTGCAACTTTTATCCGCTCCATCCAGTCTATTAAATTTGTCGGGAAGCTAG 2940
Db 2880 AAGTGTCTCTGCAACTTTTATCCGCTCCATCCAGTCTATTAAATTTGTCGGGAAGCTAG 2939
QY 2941 AGTAAAGTATTCGCCAGTTAAATAGTTTGGCGAAAGTTGTCATTTGCTACAGGCAATCGT 3000
Db 2940 AGTAAAGTATTCGCCAGTTAAATAGTTTGGCGAAAGTTGTCATTTGCTACAGGCAATCGT 2999
QY 3001 GGTGTCAAGCTCGCTGTTGGTATGCTTCAATTCAGCTCGGTTCCCAACGATCAAGGCG 3060
Db 3000 GGTGTCAAGCTCGCTGTTGGTATGCTTCAATTCAGCTCGGTTCCCAACGATCAAGGCG 3059
```

```
QY 3061 AGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCTTCGTCCTCCGATCGT 3120
Db 3060 AGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCTTCGTCCTCCGATCGT 3119
QY 3121 TGTCAAGAGTAAGTTGCGCGCAGTGTATCAGCTCATGGTTATGCGCAGCACTGCATAATTC 3180
Db 3120 TGTCAAGAGTAAGTTGCGCGCAGTGTATCAGCTCATGGTTATGCGCAGCACTGCATAATTC 3179
QY 3181 TCTTACTGTCAATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGTGAGTACTCAACCAAGTC 3240
Db 3180 TCTTACTGTCAATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGTGAGTACTCAACCAAGTC 3239
QY 3241 ATTCTGAGATAGTGTATGCGGCGACCGAGTTGCTCTTGGCCCGGCTCAATACGGATAA 3300
Db 3240 ATTCTGAGATAGTGTATGCGGCGACCGAGTTGCTCTTGGCCCGGCTCAATACGGATAA 3299
QY 3301 TACCGCGCCACATAGCAGAACTTTTAAAGTGTCTCATCATTTGGAAAAAGTTCTTCGGGGCG 3360
Db 3300 TACCGCGCCACATAGCAGAACTTTTAAAGTGTCTCATCATTTGGAAAAAGTTCTTCGGGGCG 3359
QY 3361 AAAAATCTCAAGGATCTTACCGCTGTGTGAGATCCAGTTGATGTAAACCACTCGTGCACC 3420
Db 3360 AAAAATCTCAAGGATCTTACCGCTGTGTGAGATCCAGTTGATGTAAACCACTCGTGCACC 3419
QY 3421 CAACTGATCTTACGATCTTTTACTTTTACAGCGGTTTCTGGGTGAGCAAAAAACAGGAAG 3480
Db 3420 CAACTGATCTTACGATCTTTTACTTTTACAGCGGTTTCTGGGTGAGCAAAAAACAGGAAG 3479
QY 3481 GCAAAATGCGCAAAAAAGGAATAGGCGCAGCAGCAAAATGTTGAATACTCATCTCTT 3540
Db 3480 GCAAAATGCGCAAAAAAGGAATAGGCGCAGCAGCAAAATGTTGAATACTCATCTCTT 3539
QY 3541 CCTTTTCAATATATTGTAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATT 3600
Db 3540 CCTTTTCAATATATTGTAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATT 3599
QY 3601 TGAATGTTATAGAAAAATAAACAATAGGGGTTCCGCGCATTTTCCCGAAAAAGTGCC 3660
Db 3600 TGAATGTTATAGAAAAATAAACAATAGGGGTTCCGCGCATTTTCCCGAAAAAGTGCC 3659
QY 3661 AC 3662
Db 3660 AC 3661

RESULT 2
ADM68463/C
ID ADM68463 standard; DNA; 3637 BP.
XX
AC ADM68463;
XX
XX
DT 03-JUN-2004 (first entry)
XX
DE Variant green fluorescent protein, GFP, gene construct.
XX
KW ds; mismatch endonuclease; endonuclease; gene shuffling technology;
KW single nucleotide polymorphism; cancer susceptibility;
KW sequence variation redistribution; GFP; green fluorescent protein; gene.
XX
OS Aequorea victoria.
OS Synthetic.
XX
XX US2003157682-A1.
XX
XX 21-AUG-2003.
XX
XX 31-JAN-2003; 2003US-00356708.
XX
XX 01-FEB-2002; 2002US-0353722P.
XX
XX 14-MAR-2002; 2002US-00098155.
XX
XX 01-AUG-2002; 2002US-00211079.
XX
XX (PADG/) PADGETT H S.
```


QY	1620	TCACTTAATGCGTGGTGGCTCACTGCGCGCTTTCAGTCCGGAAACCTGCTCGTGCAGC	1679
Db	2043	TCACTTAATGCGTGGTGGCTCACTGCGCGCTTTCAGTCCGGAAACCTGCTCGTGCAGC	1984
QY	1680	TGCATTAATGAATCGGCCAACGCGCGGGAGAGCGGTTTGGCTATTTGGCGCTCTTCCG	1739
Db	1983	TGCATTAATGAATCGGCCAACGCGCGGGAGAGCGGTTTGGCTATTTGGCGCTCTTCCG	1924
QY	1740	CTTCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTC	1799
Db	1923	CTTCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTC	1864
QY	1800	ACTCAAAAGCGGTAATACGGTTATCCACAGAATCAGGGGATTAACGAGAAAGAAATGT	1859
Db	1863	ACTCAAAAGCGGTAATACGGTTATCCACAGAATCAGGGGATTAACGAGAAAGAAATGT	1804
QY	1860	GAGCAAAAGCCAGCAAAAGCCAGAAACCGTAAAGAGCCGCGTTCCTGGCGCTTTTTC	1919
Db	1803	GAGCAAAAGCCAGCAAAAGCCAGAAACCGTAAAGAGCCGCGTTCCTGGCGCTTTTTC	1744
QY	1920	ATAGGCTCCGCCCTCGAGCATCACAAAATCGACCTCAAGTCAGAGGTGGCGAA	1979
Db	1743	ATAGGCTCCGCCCTCGAGCATCACAAAATCGACCTCAAGTCAGAGGTGGCGAA	1684
QY	1980	ACCGACAGGACTATTAAGATACAGCGGTTTCCCTCGAAGCTCCCTCGTGCCTCTC	2039
Db	1683	ACCGACAGGACTATTAAGATACAGCGGTTTCCCTCGAAGCTCCCTCGTGCCTCTC	1624
QY	2040	CTGTTCCGACCTGCGCTTACCGGATACCTGTCGCGCTTCTCCCTCGGAGCGGTGG	2099
Db	1623	CTGTTCCGACCTGCGCTTACCGGATACCTGTCGCGCTTCTCCCTCGGAGCGGTGG	1564
QY	2100	CGCTTCTCATAGCTCACGCTGAGGTATCTCAGTTCCGTTGAGTTCGTTCCCAAGC	2159
Db	1563	CGCTTCTCATAGCTCACGCTGAGGTATCTCAGTTCCGTTGAGTTCGTTCCCAAGC	1504
QY	2160	TGGCTGTGTGCAAGAACCCCGTTTACGCCGAGCTGCGCTTATCCGTTAACTATC	2219
Db	1503	TGGCTGTGTGCAAGAACCCCGTTTACGCCGAGCTGCGCTTATCCGTTAACTATC	1444
QY	2220	GTCTTGAGTCCAAACCGGTAAAGCAGCTTATCGCCACTGGCAGCAGCCACTGGTAA	2279
Db	1443	GTCTTGAGTCCAAACCGGTAAAGCAGCTTATCGCCACTGGCAGCAGCCACTGGTAA	1384
QY	2280	GGATTAGCAGAGGAGGTATGTAGGCGGTGTACAGAGTTCTTGAAGTGGTGGCTTAACT	2339
Db	1383	GGATTAGCAGAGGAGGTATGTAGGCGGTGTACAGAGTTCTTGAAGTGGTGGCTTAACT	1324
QY	2340	ACGGCTTACTAGAAAGCAGTATTTGGTATCTGCGCTCTGCTGAAAGCCAGTTACCTTCG	2399
Db	1323	ACGGCTTACTAGAAAGCAGTATTTGGTATCTGCGCTCTGCTGAAAGCCAGTTACCTTCG	1264
QY	2400	GAAAGAGTTGGTAGCTCTTGATCCGGCAACAAACCCAGCTGCTAGGCGGTGTTTTT	2459
Db	1263	GAAAGAGTTGGTAGCTCTTGATCCGGCAACAAACCCAGCTGCTAGGCGGTGTTTTT	1204
QY	2460	TTGTTTGCAGCAGCAGATTTACCGCAGAAAGAAAGGATCTCAAGAGATCTTTGATCT	2519
Db	1203	TTGTTTGCAGCAGCAGATTTACCGCAGAAAGAAAGGATCTCAAGAGATCTTTGATCT	1144
QY	2520	TTTCTACGGGCTGTAGCTGTAGTGAACGAAACCTCAGCTTAAGGGAATTTGGTCATGA	2579
Db	1143	TTTCTACGGGCTGTAGCTGTAGTGAACGAAACCTCAGCTTAAGGGAATTTGGTCATGA	1084
QY	2580	GATTATCAAAAGGATCTTCACTAGATCTTTTAAATTAATAATGAAGTTTAAATCAA	2639
Db	1083	GATTATCAAAAGGATCTTCACTAGATCTTTTAAATTAATAATGAAGTTTAAATCAA	1024
QY	2640	TCTAAAGTATATATAGTAAACTTGGTCTGACAGTTTACCAGTCTTAAATCAGTGAGGCAC	2699
Db	1023	TCTAAAGTATATATAGTAAACTTGGTCTGACAGTTTACCAGTCTTAAATCAGTGAGGCAC	964

RESULT 3

ADP26616/c

ID ADP26616 standard; DNA; 3637 BP.

XX

QY	2700	CTATCTCAGCGATCTCTCTATTTTCTGTTTCACTCATAGTTGCTCTGATCTCCCGCTGTTAGA	2759
Db	963	CTATCTCAGCGATCTCTCTATTTTCTGTTTCACTCATAGTTGCTCTGATCTCCCGCTGTTAGA	904
QY	2760	TAACTACGATACGGGAGGCTTACCATCTGGCCCCAGTCTGCAATGATACCCGAGACC	2819
Db	903	TAACTACGATACGGGAGGCTTACCATCTGGCCCCAGTCTGCAATGATACCCGAGACC	844
QY	2820	CACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCAGCCGGAAGCCGAGCGCA	2879
Db	843	CACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCAGCCGGAAGCCGAGCGCA	784
QY	2880	GAAAGTGGTCTGCAACTTTTATCCGCTCCATCCAGTCTATTAATTTGTTGCGGGAAGCTA	2939
Db	783	GAAAGTGGTCTGCAACTTTTATCCGCTCCATCCAGTCTATTAATTTGTTGCGGGAAGCTA	724
QY	2940	GAGTAAGTAGTTTCCGCAAGTTTATAGTTTGGCAACGTTGTTGCATTTGCTACAGGCAATC	2999
Db	723	GAGTAAGTAGTTTCCGCAAGTTTATAGTTTGGCAACGTTGTTGCATTTGCTACAGGCAATC	664
QY	3000	TGGTGTACGCTCGTCTGTTTGGTATGGCTTCAATTCAGCTCCGTTCCCAACGATCAAGGC	3059
Db	663	TGGTGTACGCTCGTCTGTTTGGTATGGCTTCAATTCAGCTCCGTTCCCAACGATCAAGGC	604
QY	3060	GAGTTACATGATCCCCCATGTTTGTGCAAAAAGCGGTAGCTCTTCGCTCCCGATCG	3119
Db	603	GAGTTACATGATCCCCCATGTTTGTGCAAAAAGCGGTAGCTCTTCGCTCCCGATCG	544
QY	3120	TTGTCAAGAGTAAAGTTGGCGGAGTGTATCACTCATGTTATGGCAGCACTGCATAATT	3179
Db	543	TTGTCAAGAGTAAAGTTGGCGGAGTGTATCACTCATGTTATGGCAGCACTGCATAATT	484
QY	3180	CTCTTACTCTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGT	3239
Db	483	CTCTTACTCTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGT	424
QY	3240	CAATCTGAGAAATAGTGTATGCGCGGACCGAGTTGCTCTTCCCGCGGTCAATACGGGATA	3299
Db	423	CAATCTGAGAAATAGTGTATGCGCGGACCGAGTTGCTCTTCCCGCGGTCAATACGGGATA	364
QY	3300	ATACCGCGCCACATAGCAGAACTTTTAAAGTGTCTCATCTGGAATAACGTTCTTCGGGC	3359
Db	363	ATACCGCGCCACATAGCAGAACTTTTAAAGTGTCTCATCTGGAATAACGTTCTTCGGGC	304
QY	3360	GAAAACTCTCAAGGATCTTACCGCTGTTCAGATGCCAGTGTAAACCCACTCGTGCAC	3419
Db	303	GAAAACTCTCAAGGATCTTACCGCTGTTCAGATGCCAGTGTAAACCCACTCGTGCAC	244
QY	3420	CCAACTGTCTTCAGCATCTTTTACCTTTCACAGCGTTTCTGGGTGAGCAAAACAGGAA	3479
Db	243	CCAACTGTCTTCAGCATCTTTTACCTTTCACAGCGTTTCTGGGTGAGCAAAACAGGAA	184
QY	3480	GGCAAAATGCCCAAAAAGGGAATAAGGCGCAGACGGAATAATGTTCAATCTCATCTCT	3539
Db	183	GGCAAAATGCCCAAAAAGGGAATAAGGCGCAGACGGAATAATGTTCAATCTCATCTCT	124
QY	3540	TCTTTTCAATTAATTTGAAGCATTTATCAGGTTTATTTGCTCATGAGCGGATACATAT	3599
Db	123	TCTTTTCAATTAATTTGAAGCATTTATCAGGTTTATTTGCTCATGAGCGGATACATAT	64
QY	3600	TTGAATGTTATAGAAAAATAAACAAATAGGGGTTCCCGCGCAGCATTTCCCGAAAAAGTGC	3659
Db	63	TTGAATGTTATAGAAAAATAAACAAATAGGGGTTCCCGCGCAGCATTTCCCGAAAAAGTGC	4
QY	3660	CAC 3662	
Db		3 CAC 1	

AC ADP26616;
XX
DT 26-AUG-2004 (first entry)
XX
DE Green fluorescent protein (GFP) plasmid DNA #3.
XX
KW Sequence variation; heteroduplex; transcription; DNA integration;
KW ribozyme expression; gene; ds; green fluorescent protein; GFP.
XX
XX Aequorea victoria.
OS Synthetic.
XX
PN US2004110130-A1.
XX
XX 10-JUN-2004.
XX
XX 25-OCT-2002; 2002US-00280913.
XX
XX 02-FEB-2001; 2001US-0266386P.
PR 14-FEB-2001; 2001US-0266785P.
PR 01-FEB-2002; 2002US-0006639P.
PR 08-AUG-2002; 2002US-0402342P.
XX
XX (LARG-) LARGE SCALE BIOLOGY CORP.
XX
XX Padgett HS, Lindbo JA, Fitzmaurice WP;
PI WPI; 2004-440326/41.
DR
XX
XX Redistributing sequence variations between non-identical polynucleotide
PT sequences, useful for generating improved polynucleotide having a desired
PT characteristic, comprises making a heteroduplex and introducing a nick.
XX
XX Example 14; SEQ ID NO 32; 75pp; English.
PS
XX
XX The invention relates to an in vitro method of redistributing sequence
CC variations between non-identical polynucleotide sequences, comprising
CC making a heteroduplex polynucleotide from two non-identical
CC polynucleotides, introducing a nick in the second strand at or near a
CC base pair mismatch site, removing the mismatched base(s) from the
CC mismatch site where the nick occurred and using the first strand as a
CC template to replace the removed base(s) with bases that complement the
CC base(s) in the first strand. The invention also relates to an in vitro
CC method of making a population of sequence variants from a heteroduplex
CC polynucleotide sequence, obtaining a polynucleotide sequence encoding a
CC desired functional property and identifying a reassorted DNA molecule
CC encoding a protein with a desired functional property. The method is
CC useful for generating an improved polynucleotide sequence or a population
CC of improved polynucleotide sequences possessing at least one desired
CC phenotypic characteristic (e.g., promotes transcription of linked
CC polynucleotides), where such polynucleotides are useful for expression
CC from a plant, animal, fungal, yeast, or bacterial expression vector, for
CC integration to form a transgenic plant, animal or microorganism, and for
CC expression of a ribozyme. This sequence represents DNA used in the scope
CC of the invention.
XX
SQ Sequence 3637 BP; 942 A; 891 C; 870 G; 934 T; 0 U; 0 Other;

Query Match 72.8%; Score 2665.4; DB 12; Length 3637;
Best Local Similarity 83.7%; Pred. No. 0;
Matches 3065; Conservative 0; Mismatches 571; Indels 27; Gaps 3;

QY 1 CTAATTCGTACGCGTAAATATTTGTTAAATTCGCGTTAAATTTTGTGTTAAATCAGCTC 60
DB |||||||
DB 3637 CTAATTCGTACGCGTAAATATTTGTTAAATTCGCGTTAAATTTTGTGTTAAATCAGCTC 3578

QY 61 ATTTTTTAAACCAATAGCGCGAAATCGCAAAATCCCTTATAAATCAAAGAATAGACCGA 120
DB |||||||
DB 3577 ATTTTTTAAACCAATAGCGCGAAATCGCAAAATCCCTTATAAATCAAAGAATAGACCGA 3518

QY 121 GATAGGTTGAGTGTGTTCCAGTTTGGNACAAGAGTCCACTATTAAAGAACTGGACTC 180
DB |||||||
DB 3517 GATAGGTTGAGTGTGTTCCAGTTTGGNACAAGAGTCCACTATTAAAGAACTGGACTC 3458

QY 181 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 240
DB |||||||
DB 3457 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 3398

QY 241 CTAATCAAGTTTTTTGGGGTCGAGTGCCTGTAAGACATAAATCGGAACCCCTAAAGGGAG 300
DB |||||||
DB 3397 CTAATCAAGTTTTTTGGGGTCGAGTGCCTGTAAGACATAAATCGGAACCCCTAAAGGGAG 3338

QY 301 CCCCCGATTTAGAGCTTCAOCCGGGAAGCCGCGCAACCTGGCGAGAAAGGAGGGAAGAA 360
DB |||||||
DB 3337 CCCCCGATTTAGAGCTTCAOCCGGGAAGCCGCGCAACCTGGCGAGAAAGGAGGGAAGAA 3278

QY 361 AGCGAAAGGAGCGGCGCTTAGGGCGCTGCGCAAGTGTAGCGGTACGCTGCGGTAAACAC 420
DB |||||||
DB 3277 AGCGAAAGGAGCGGCGCTTAGGGCGCTGCGCAAGTGTAGCGGTACGCTGCGGTAAACAC 3218

QY 421 CACACCCCGCGCTTAATTCGCGCCCTACAGGGCGCTCCCATTCGCCANTCAGGCTGCG 480
DB |||||||
DB 3217 CACACCCCGCGCTTAATTCGCGCCCTACAGGGCGCTCCCATTCGCCANTCAGGCTGCG 3158

QY 481 CAACTGTTGGGNAAGGCGATCGGTGCGGGCCCTCTCGCTATTACGCCAGCTCGCGAAAGG 540
DB |||||||
DB 3157 CAACTGTTGGGNAAGGCGATCGGTGCGGGCCCTCTCGCTATTACGCCAGCTCGCGAAAGG 3098

QY 541 GGGATGCTGTCGAAGCGATTAAAGTTGGGTAAACGCCAGGGTTTTCCCACTACGACGTTG 600
DB |||||||
DB 3097 GGGATGCTGTCGAAGCGATTAAAGTTGGGTAAACGCCAGGGTTTTCCCACTACGACGTTG 3038

QY 601 TAAAAAGCGGCGAGTGAAGCGGCTCGTTCAATTCACGTTTTTTGAAACCGTGGAGACGG 660
DB |||||||
DB 3037 TAAAAAGCGGCGAGTGAAGCGGCTCGTTCAATTCACGTTTTTTGAAACCGTGGAGTCCA 2978

QY 661 GCAGACTCGCGTGCAGAAATGTTTTACAGGTGATGGAGCAGATGAAGATGCTCGACAC 720
DB |||||||
DB 2977 CGCGGTGGCGGCCGCTCTAGAACTAGTGGAATCCCGGGCTGCGAGAAATCTTATTGTT 2918

QY 721 GCTGCAGAACACCGCAGCTAGATTAAACCTAGAAAGATAATCATATTGTGACGTAGTTAA 780
DB |||||||
DB 2917 ATAGTTTCATCCATGCCATGTGTAATCCCGCAGCAGAGTTTACAACTCAAGAAGCAGCATG 2858

QY 781 AGATAATC-ATGCGTAAATTTGACGCATGGGATCTGTAAATAGCAGCTCATATAGGCGNA 839
DB |||||||
DB 2857 GGTACACGCTTTTCGTTGGGATCTTTTCGAAAGGGCAGATTGTGTGACACAGGTAATGTTG 2798

QY 840 TTGGGTACGGGCCCCCTCGAGGTGCGAGGTATCGATTAAGCTTGAATCGAATTCCTG 899
DB |||||||
DB 2797 CTGGTAAAGGACACAGGCGCATCGCCAAATGGAGTATTTTGTGATAATGGTCTGTAGTT 2738

QY 900 CAGCCCGGGGATCCACTAGTTCTAGAGCGGCGCCACCGCGGTGGAGCTCCAGCTTTTG 959
DB |||||||
DB 2737 GAACGGATCCATCTTCAATGTTGTGCGAAATTTTGAAGTAGCTTTGATTCATCTTTT 2678

QY 960 TTCCTTTAGTGAGGGTTAAATTAGATCCCATCGGTCAATTTTACGAGACTATCTTTCTA 1019
DB |||||||
DB 2677 GTTGTCTCGCGTGAATATATATTGTGTAGTTTAAAGTTGTACTCGAGTTTGTGTCCGA 2618

QY 1020 GGGTTAATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCAATGC 1079
DB |||||||
DB 2617 GAATGTTTCCATCTTCTTTAAATCAATFACCTTTTAACTCGATAC-----GATTAAACAG 2563

QY 1080 CCAAGCTCGGCTATCTGGGCATCGGGGAGGAAGCCCGGTGCGCTTTTCCCGCGAGGTT 1139
DB |||||||
DB 2562 GGTATCACCTTCAACTTGAAGTTGACGACGGGTCTTGTAGTTCCCGTCACTCTTTGAAGA 2503

QY 1140 GAAGCGGATGAAGAAGGTTTGGCCGAGGATGATGCTGCTGATGAGCTTGAGCGGAGAA 1199
DB |||||||
DB 2502 TATAGTGGCTTCTGTACATAACCTTCGGGCATGGCACTCTTTGAAAAAGATCATGCCGTT 2443

QY 1200 GCGAGTTTACCATCATGATTCGGGAAGGTGGGCCATGCACGCTTTTAAACGTTGAACGTG 1259
DB |||||||
DB 2442 CATATGTCGGATTAACGAGAAAGCAATTAACACCATGAGAGAAAGTAGTGACAAAGTGT 2383

QY 1260 TTGTTTCCAGCCACCTGGGATACCACTTGGTGGCGGCTTTTCGGGACACAGTTCCGGATG 1319
Db 2382 TGCCCATAGAAAGGAGTATTTTCAGTAGTGCAATAAATTTAAGTGTAAAGCTTTCCGTA 2323
QY 1320 GTGAGCCGGAAGCGCATCAGCAACCCGAAACAATACCGGGACAGCCGGAACCTGCCGTGC 1379
Db 2322 TGTAGCATCACCTTACCTCTCCACTGACAGAAATTTGTGCCCATTAACATCACCATC 2263
QY 1380 GGTGTCCAGATTAATGACAGCGGTGGCGCTGGGATATTAAGTACGAGGAGCGGGTAT 1439
Db 2262 TAAATCAACAAGAAAT-----TGGGACAACTCCAGTGAAAAAGTTC 2224
QY 1440 CTTGGCTGGATCCGACAGAAATGGAATGATGATACCCCGGTGAGTTACCGCGGGCGCGCT 1499
Db 2223 TTCCTCTTACTCATCGGTACCCAGCTTTTGTCCCTTTAGTGAGGGTAAATTTGCCGCT 2164
QY 1500 TGGCGTAATCATGGTCATAGCTGTGTTTCTGTGTGAATTTGTTATCCGCTCACAAATTCAC 1559
Db 2163 TGGCGTAATCATGGTCATAGCTGTGTTTCTGTGTGAATTTGTTATCCGCTCACAAATTCAC 2104
QY 1560 ACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTTGGGGTGCCTAATGAGTGAGCTAAC 1619
Db 2103 ACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTTGGGGTGCCTAATGAGTGAGCTAAC 2044
QY 1620 TCACATTAATTTGGTTCGCTCACTGCCGCTCACTGCCGCTTTCAGTCCGGAAACCTGTGTCGACG 1679
Db 2043 TCACATTAATTTGGTTCGCTCACTGCCGCTTTCAGTCCGGAAACCTGTGTCGACG 1984
QY 1680 TGCATTAATTAATTCGGCCAAAGCGGGGAGAGCGGTTTGGCTATTTGGCGCTTTTCG 1739
Db 1983 TGCATTAATTAATTCGGCCAAAGCGGGGAGAGCGGTTTGGCTATTTGGCGCTTTTCG 1924
QY 1740 CTTTCCTCGCTACTGACTCGTGCCTCGGTGCTTCCGCTGCGGCGAGCGGTATCAGCTC 1799
Db 1923 CTTTCCTCGCTACTGACTCGTGCCTCGGTGCTTCCGCTGCGGCGAGCGGTATCAGCTC 1864
QY 1800 ACTCAAGCGGTAATACGGTTATCCACAGATCAGGGATACCGGGAAGAACATGT 1859
Db 1863 ACTCAAGCGGTAATACGGTTATCCACAGATCAGGGATACCGGGAAGAACATGT 1804
QY 1860 GAGCAAAAGCGCAGCAAAAGCGCAGAACCGTAAAGAGCGCGTTGCTGGGCTTTTCC 1919
Db 1803 GAGCAAAAGCGCAGCAAAAGCGCAGAACCGTAAAGAGCGCGTTGCTGGGCTTTTCC 1744
QY 1920 ATAGGCTCCGCCCTCTGACGAGCATCACAAAAATCAGCCTCAAGTCAGAGGTGGCGAA 1979
Db 1743 ATAGGCTCCGCCCTCTGACGAGCATCACAAAAATCAGCCTCAAGTCAGAGGTGGCGAA 1684
QY 1980 ACCGACAGGACTATAAGATACAGCGGTTCCTCCCTGGAAGCTCCCTCGTGGCTCTC 2039
Db 1683 ACCGACAGGACTATAAGATACAGCGGTTCCTCCCTGGAAGCTCCCTCGTGGCTCTC 1624
QY 2040 CTGTTCCGACCTCGCGCTTACCGGATACCTGTCCGCTTCTCCCTCGGGAAGCGTG 2099
Db 1623 CTGTTCCGACCTCGCGCTTACCGGATACCTGTCCGCTTCTCCCTCGGGAAGCGTG 1564
QY 2100 CGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTTCGCTGTAGGTGCTTCGCTCAAGC 2159
Db 1563 CGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTTCGCTGTAGGTGCTTCGCTCAAGC 1504
QY 2160 TGGGTGTGTGACGAAACCCCGTTTCAGCCGACCGCTGGCGCTTATCCGTTAATATC 2219
Db 1503 TGGGTGTGTGACGAAACCCCGTTTCAGCCGACCGCTGGCGCTTATCCGTTAATATC 1444
QY 2220 GTCTTGAAGTCCAAACCGGTAAAGACGACTTATCCGCTGCGACGACCACTGGTAACA 2279
Db 1443 GTCTTGAAGTCCAAACCGGTAAAGACGACTTATCCGCTGCGACGACCACTGGTAACA 1384
QY 2280 GGATTAGCAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTTAAT 2339
Db 1383 GGATTAGCAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTTAAT 1324
QY 2340 ACGGCTACACTAGAAAGGACAGTATTTGGTATCTGGCGCTCTGCTGAAGCCAGTTACCTCG 2399

Db 1323 ACGGCTACACTAGAAAGGACAGTATTTGGTATCTGCCCTCTGCTGAAGCCAGTTACCTTCG 1264
QY 2400 GAAAAAGAGTTGGTAGCTCTTGTATCCGCAAAACAAACCCGCTGGTAGCGGTGGTTTTT 2459
Db 1263 GAAAAAGAGTTGGTAGCTCTTGTATCCGCAAAACAAACCCGCTGGTAGCGGTGGTTTTT 1204
QY 2460 TTGTTTGAAGCGAGATATACCGCAGAAAAAAGGATCTCAAGAGATCTTGTGATCT 2519
Db 1203 TTGTTTGAAGCAGCAGATATACCGCAGAAAAAAGGATCTCAAGAGATCTTGTGATCT 1144
QY 2520 TTTCTACGGGCTGTGACGCTCAGTGGAACGAAACCTCAGTTAAGGATTTTGGTCAATGA 2579
Db 1143 TTTCTACGGGCTGTGACGCTCAGTGGAACGAAACCTCAGTTAAGGATTTTGGTCAATGA 1084
QY 2580 GATTATCAAAAAGGATCTTCCACCTAGATCTTTTAAATTAATAATGAAGTTTAAATCAA 2639
Db 1083 GATTATCAAAAAGGATCTTCCACCTAGATCTTTTAAATTAATAATGAAGTTTAAATCAA 1024
QY 2640 TCTAAAGTATATAGATGATTAACCTTGGTCTGACAGTTTCAATGCTTAAATCAGTGAGGCAC 2599
Db 1023 TCTAAAGTATATAGATGATTAACCTTGGTCTGACAGTTTCAATGCTTAAATCAGTGAGGCAC 964
QY 2700 CTATCTCAGCGATCTGTCTATTTTGGTTCATTCATAGTTCGCTGACTCCCGCTCGTGTAGA 2759
Db 963 CTATCTCAGCGATCTGTCTATTTTGGTTCATTCATAGTTCGCTGACTCCCGCTCGTGTAGA 904
QY 2760 TAACTACGATACGGGAGGCTTACCATCTGGGCCAGTGTGCAATGATACCCGAGAC 2819
Db 903 TAACTACGATACGGGAGGCTTACCATCTGGGCCAGTGTGCAATGATACCCGAGAC 844
QY 2820 CACGCTCACCGGCTCCAGATTTATCAGCAATTAACCCAGCAGCCGGAAGGCGCAGCGCA 2879
Db 843 CACGCTCACCGGCTCCAGATTTATCAGCAATTAACCCAGCAGCCGGAAGGCGCAGCGCA 784
QY 2880 GAAGTGGTCTCCGACTTTATCGCTCCATCCAGTCTATTAATTCCTGCGCGGAGCTA 2939
Db 783 GAAGTGGTCTCCGACTTTATCGCTCCATCCAGTCTATTAATTCCTGCGCGGAGCTA 724
QY 2940 GAGTAAGTAGTTTCGCGCAGTAAATAGTTTCGCAACCGTTGTGCAATGCTACAGGATCG 2999
Db 723 GAGTAAGTAGTTTCGCGCAGTAAATAGTTTCGCAACCGTTGTGCAATGCTACAGGATCG 664
QY 3000 TGGTGTCAAGCTCGTGTGTTGGTATGGCTTCAATCAGCTCCGTTCCCAACGATCAAGGC 3059
Db 663 TGGTGTCAAGCTCGTGTGTTGGTATGGCTTCAATCAGCTCCGTTCCCAACGATCAAGGC 604
QY 3060 GAGTTACATGATCCCGATTTGTCGCAAAAGCGGTAGCTCTTCCGCTCCGATCG 3119
Db 603 GAGTTACATGATCCCGATTTGTCGCAAAAGCGGTAGCTCTTCCGCTCCGATCG 544
QY 3120 TTGTCAGAAAGTAAAGTTGGCGCAGTGTATCACTCATGTGTTATGGCAGCACTGCATAAT 3179
Db 543 TTGTCAGAAAGTAAAGTTGGCGCAGTGTATCACTCATGTGTTATGGCAGCACTGCATAAT 484
QY 3180 CTCTTACTGTCTAGCCATCCGTAAGATGCTTTTGTGTACTGTGTAGTACTCAACCAAGT 3239
Db 483 CTCTTACTGTCTAGCCATCCGTAAGATGCTTTTGTGTACTGTGTAGTACTCAACCAAGT 424
QY 3240 CATTCAGAAATAGTGTATGCGGACCGAGTTGCTTTCGCGCGGCTCAATACGGGATA 3299
Db 423 CATTCAGAAATAGTGTATGCGGACCGAGTTGCTTTCGCGCGGCTCAATACGGGATA 364
QY 3300 ATACCGCGCACATAGCAGAACTTTTAAAGTGTCTCATCTGGAAGAAAGCTTCTTCGGGGC 3359
Db 363 ATACCGCGCACATAGCAGAACTTTTAAAGTGTCTCATCTGGAAGAAAGCTTCTTCGGGGC 304
QY 3360 GAAAACTCTCAAGGATCTTACCGCTGTGTAGATCCAGTTTCAAGTAAACCCACTCGTGAC 3419
Db 303 GAAAACTCTCAAGGATCTTACCGCTGTGTAGATCCAGTTTCAAGTAAACCCACTCGTGAC 244
QY 3420 CCAACTGATCTTCAGGATCTTTTACTTTTCAACGAGCTTCTGCGTGAGCAAAACAGGAA 3479

Db 2737 GAAAGGATCCATCTTCAATGTTGGCGGAATTTTGAAGTTAGCTTTGATTCATTTCTTTT 2678
Qy 960 TTCCCTTTTAGTGAGGGTTAAATAGATCCCATGCGTCAAAATTTTACGAGACTATCTTTCTA 1019
Db 2677 GTTTGTCTGCGGTGATACATTTGTGTAGTTAAAGTTGTACTCGAGTTGTGTCCGA 2618
Qy 1020 GGGTTAAATCTAGCTGATCAGGATCAATATGTCGGGTCTTTTTCGGGCTCAGTCAATCGC 1079
Db 2617 GAAATGTTTCCATCTTCTTTTAAATCAATACCTTTTAACTCGATAC-----GATTAACAAG 2563
Qy 1080 CCAAGCTGGCGTATCTGGGCAATCGGGAGGAAGACCGCGCTTTTCCCGGAGGTT 1139
Db 2562 GGTATCACCTTCAAACCTTGACTTCACACGCGTCTGTAGTTCCCGTCACTTTTGAAGA 2503
Qy 1140 GAAGCGGCATGGAAGAGGTTTCCGAGGATGACTGCTGCTGCAATTAACGTTGAGCGAAAA 1199
Db 2502 TATAGTGGCTTCTCTGTACATAACCTTCGGGATGGCACTCTTGAAAAAGTTCATGCGGTTT 2443
Qy 1200 CGCAGTTTACCATGATGATTCGGGAAGGTGGGCCATGACGCCCTTTTAAAGGTGAACGTG 1259
Db 2442 CATATGATCCGGATAACGAGAAAGCAITTGAAACACCATGAGAGAAAAGTAGTGAAGTGT 2383
Qy 1260 TTGTTTCAGGCCACCTGGGATACAGTTCGTGCGGCTTTTCCGGACACAGTTCCGGATG 1319
Db 2382 TGGCCATGGAACAGGTAGTTTCCAGTAGTGCAAAATTTTAAAGTGAAGCTTTCCGTA 2323
Qy 1320 GTCAGCGGGAAGCGCATGACGAACCCGAAACCAATPACCGGCGACAGCGGAACCTGCCGTGC 1379
Db 2322 TGTAGCATCACCTTTCACCTCTCCACTGACAGAAATTTTGTGCCATTAAACATCACCATC 2263
Qy 1380 GGTGTCCAGATTAATGACAGCGGTGGCGCTGGGATATTAAGTCAAGAGAGCGGGTAT 1439
Db 2262 TAAATCAACAAGAAAT-----TGGGACAACTCCAGTGAAGAAAGTTTC 2224
Qy 1440 CCTGGCTGATGCCGAGAAATGGACATGGATACCCCGTGGAGTTACCCGGGGCGCGCT 1499
Db 2223 TTCTCTCTTACTCATCGGTACCCAGCTTTTGTTCCTTTAGTGAGGGTTAAATGGCGCGCT 2164
Qy 1500 TGGCGTAATCATGGTCTATAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCACAATTTCCAC 1559
Db 2163 TGGGTAAATCATGGTCTATAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCACAATTTCCAC 2104
Qy 1560 ACAACATACGAGCGGAAGCATAAAGTGTAAAGCTTGGGGTGCTTAATCAGTGAGCTAAC 1619
Db 2103 ACAACATACGAGCGGAAGCATAAAGTGTAAAGCTTGGGGTGCTTAATCAGTGAGCTAAC 2044
Qy 1620 TCACATTAAATGGGTGGCTCACTGCCGCTTTCAGTCCGGAACCTGTGTCGCAGC 1679
Db 2043 TCACATTAAATGGGTGGCTCACTGCCGCTTTCAGTCCGGAACCTGTGTCGCAGC 1984
Qy 1680 TGCATTAAATGAATCCGCCAAACGCGGGGAGAGCGGTTTGGGTATTTGGGCGCTCTTCCG 1739
Db 1983 TGCATTAAATGAATCCGCCAAACGCGGGGAGAGCGGTTTGGGTATTTGGGCGCTCTTCCG 1924
Qy 1740 CTTCTCGCTCACTGACTCGCTGCGCTCGGTGCTTGGGCTGGGCGAGCGGTATCAGCTC 1799
Db 1923 CTTCTCGCTCACTGACTCGCTGCGCTCGGTGCTTGGGCTGGGCGAGCGGTATCAGCTC 1864
Qy 1800 ACTCAAGGCGGTATACGGTTATCCACAGATCAGGGGTAACCGAGGAAGACATGT 1859
Db 1863 ACTCAAGGCGGTATACGGTTATCCACAGATCAGGGGTAACCGAGGAAGACATGT 1804
Qy 1860 GAGCAAAAGGCGAGCAAAAGGCGAGGAACCGTAAAAAGGCGGTTTGGCTTTTCC 1919
Db 1803 GAGCAAAAGGCGAGCAAAAGGCGAGGAACCGTAAAAAGGCGGTTTGGCTTTTCC 1744
Qy 1920 ATAGGCTCCGCCCTCTGACGAGCATCACAAAAATCAGCTCAAGTCAAGGTGGCGAA 1979
Db 1743 ATAGGCTCCGCCCTCTGACGAGCATCACAAAAATCAGCTCAAGTCAAGGTGGCGAA 1684
Qy 1980 ACCGACGAGACTATAAAGATACAGGCGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTC 2039
|||||

Db 1683 ACCGACGAGACTATAAAGATACAGGCGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTC 1624
Qy 2040 CTGTTCCGACCTGCGCTTACCGGATACCTGTGCGCTTTCTCCCTCGGGAAGCGTGG 2099
Db 1623 CTGTTCCGACCTGCGCTTACCGGATACCTGTGCGCTTTCTCCCTCGGGAAGCGTGG 1564
Qy 2100 CGCTTCTCATAGCTCAGCTGTAGGTATCTCAGTTTGGGTGAGGTGCTTCCGCTCCAGC 2159
Db 1563 CGCTTCTCATAGCTCAGCTGTAGGTATCTCAGTTTGGGTGAGGTGCTTCCGCTCCAGC 1504
Qy 2160 TGGGCTGTGTGACGAAACCCCGCTTACGCGGACCGCTGCGCTTTATCCGGTAACTATC 2219
Db 1503 TGGGCTGTGTGACGAAACCCCGCTTACGCGGACCGCTGCGCTTTATCCGGTAACTATC 1444
Qy 2220 GTCTTGAGTCCAAACCGGTAAGACAGACTTATCGCACTGGCAGCAGGCACCTGGTAACA 2279
Db 1443 GTCTTGAGTCCAAACCGGTAAGACAGACTTATCGCACTGGCAGCAGGCACCTGGTAACA 1384
Qy 2280 GGATTAGCAGAGGAGGTATGTAGCGGTGCTACAGAGTCTTGAAGTGGTGGCCCTAACT 2339
Db 1383 GGATTAGCAGAGGAGGTATGTAGCGGTGCTACAGAGTCTTGAAGTGGTGGCCCTAACT 1324
Qy 2340 ACGCTTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAAGCCAGTTACCTTCG 2399
Db 1323 ACGCTTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAAGCCAGTTACCTTCG 1264
Qy 2400 GAAAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACCAACCGCTGCTGAGCGGTGTTTTT 2459
Db 1263 GAAAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACCAACCGCTGCTGAGCGGTGTTTTT 1204
Qy 2460 TTGTTTGAAGCAGCAGATTACCGCAGAAAAAAGGATCTCAAGAGATCTCTTGTATCT 2519
Db 1203 TTGTTTGAAGCAGCAGATTACCGCAGAAAAAAGGATCTCAAGAGATCTCTTGTATCT 1144
Qy 2520 TTCTTACGGGTCTGACGCTCAGTGGAACGAAAACTCAGTTAAAGGATTTTGGTCAATGA 2579
Db 1143 TTCTTACGGGTCTGACGCTCAGTGGAACGAAAACTCAGTTAAAGGATTTTGGTCAATGA 1084
Qy 2580 GATTATCAAAAAAGGATCTTACCTAGATCTTTTAAATTTAAATTTAAATTTAAATTTAAATCA 2639
Db 1083 GATTATCAAAAAAGGATCTTACCTAGATCTTTTAAATTTAAATTTAAATTTAAATTTAAATCA 1024
Qy 2640 TCTAAAGTATATAGGTAAACTTGGTCTGACAGATTACCAATGCTTTAATCAGTGAGGCAC 2699
Db 1023 TCTAAAGTATATAGGTAAACTTGGTCTGACAGATTACCAATGCTTTAATCAGTGAGGCAC 964
Qy 2700 CTATCTCAGCGATCTGTCTATTTCTGTTTCAATCATAGTTCGCTGACTCCCGCTGCTAGA 2759
Db 963 CTATCTCAGCGATCTGTCTATTTCTGTTTCAATCATAGTTCGCTGACTCCCGCTGCTAGA 904
Qy 2760 TAACTACGATACGGAGGCGCTTACCATCTGGCCCCCAGTCTGCAATGATACCGCGAGCC 2819
Db 903 TAACTACGATACGGAGGCGCTTACCATCTGGCCCCCAGTCTGCAATGATACCGCGAGCC 844
Qy 2820 CAGCTCACCGGCTCCAGTTTATCAGCAATAAACACAGCAGCCGGAAGGCGCGAGCGCA 2879
Db 843 CAGCTCACCGGCTCCAGTTTATCAGCAATAAACACAGCAGCCGGAAGGCGCGAGCGCA 784
Qy 2880 GAAGTGGTCTTCAACTTTTATCCGCTCCATCCAGTCTATTAATTTCTGCGGGAAGCTA 2939
Db 783 GAAGTGGTCTTCAACTTTTATCCGCTCCATCCAGTCTATTAATTTCTGCGGGAAGCTA 724
Qy 2940 GAGTAAGTAGTTCCGCGAGTTAATAGTTTTCGCAACGTTTGTGCAATTTGCTACAGGATCG 2999
Db 723 GAGTAAGTAGTTCCGCGAGTTAATAGTTTTCGCAACGTTTGTGCAATTTGCTACAGGATCG 664
Qy 3000 TGGGTGACGCTCGCTGTTTGGTATGGCTTCAATTCAGCTCCGGTTCCCAACGATCAAGGC 3059
Db 663 TGGGTGACGCTCGCTGTTTGGTATGGCTTCAATTCAGCTCCGGTTCCCAACGATCAAGGC 604
Qy 3060 GAGTTTACATGATCCCCCAATGTTTGTGCAAAAAAGCGGTAGCTTCTCGGCTCCCGATCG 3119
Db 603 GAGTTTACATGATCCCCCAATGTTTGTGCAAAAAAGCGGTAGCTTCTCGGCTCCCGATCG 544
|||||

QY 3120 TTGTGAGAGTAAGTTGGCGAGTGTTATCACTCATGCTTATGGCAGCACTGCATAATT 3179
DB |||||
543 TTGTGAGAGTAAGTTGGCGAGTGTTATCACTCATGCTTATGGCAGCACTGCATAATT 484
QY 3180 CTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAAGTACTCAACCAAGT 3239
DB |||||
483 CTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAAGTACTCAACCAAGT 424
QY 3240 CATTCTGAGAAATAGTATGGCGGACGAGTGTCTTTGCGCGGCTCAATACGGGATA 3299
DB |||||
423 CATTCTGAGAAATAGTATGGCGGACGAGTGTCTTTGCGCGGCTCAATACGGGATA 364
QY 3300 ATACCGCGCCACATAGCAGAACTTTTAAAGTGTCTATCATTTGAAAAACGTTCTTCGGGCG 3359
DB |||||
363 ATACCGCGCCACATAGCAGAACTTTTAAAGTGTCTATCATTTGAAAAACGTTCTTCGGGCG 304
QY 3360 GAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTTCGATGTAACCCACTCGTGAC 3419
DB |||||
303 GAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTTCGATGTAACCCACTCGTGAC 244
QY 3420 CCAACTGATCTTCAGCATCTTTTACTTTTCCAGCGTTTCTGGTGAGCAAAAACAGGAA 3479
DB |||||
243 CCAACTGATCTTCAGCATCTTTTACTTTTCCAGCGTTTCTGGTGAGCAAAAACAGGAA 184
QY 3480 GGCAGAAATGCCCAAAAGGAATAAGGGCGACACGGAAATGTTGAATFACTACTCTCT 3539
DB |||||
183 GGCAGAAATGCCCAAAAGGAATAAGGGCGACACGGAAATGTTGAATFACTACTACTCTCT 124
QY 3540 TCCCTTTTCAATATTAATGAAGCATTTATCAGGGTTATTGTCATGAGCGGATACATAT 3599
DB |||||
123 TCCCTTTTCAATATTAATGAAGCATTTATCAGGGTTATTGTCATGAGCGGATACATAT 64
QY 3600 TTGAATGATTTAGAAAAATAACAAATAGGGGTTCCCGGCGACATTTCCCGGAAAGTGC 3659
DB |||||
63 TTGAATGATTTAGAAAAATAACAAATAGGGGTTCCCGGCGACATTTCCCGGAAAGTGC 4
QY 3660 CAC 3662
DB |||
3 CAC 1

RESULT 5
ADSL192/c
ID ADS17192 standard; DNA; 3637 BP.
XX
AC ADS17192;
XX
DT 02-DEC-2004 (first entry)
XX
DE pBSC3BFP plasmid DNA encoding A. victoria BFP Cycle 3 gene.
XX
KW Polymerase; BFP; jellyfish; blue fluorescent protein; chimeric; gene; ds.
XX
OS Aequorea victoria.
OS Chimeric.
OS Unidentified.
XX US2004180352-A1.
XX
PD 16-SEP-2004.
XX
PF 08-AUG-2003; 2003US-00637758.
XX
PR 08-AUG-2002; 2002US-0402342P.
XX (LARG-) LARGE SCALE BIOLOGY CORP.
XX Padgett HS, Lindbo JA, Fitzmaurice WP, Vaewhongs AA;
XX WPI; 2004-667658/65.
PT In vitro method of making sequence variants from heteroduplex

PT polynucleotide, involves combining polynucleotide with agent having
PT polymerase and strand cleavage activities for sufficient time for
PT percentage of complementarity to increase.
XX Example 18; SEQ ID NO 32; 81pp; English.
XX The invention relates to an in vitro method of making sequence variants
CC from heteroduplex polynucleotide which involves combining polynucleotide
CC with an agent having polymerase and strand cleavage activities for
CC sufficient time for the percentage of complementarity to be increased
CC within the heteroduplex. The method is useful for performing in vitro
CC method of making sequence variants from one or more heteroduplex
CC polynucleotide. It is useful in increasing diversity in a population of
CC sequences and obtaining a polynucleotide encoding a desired functional
CC property. The present sequence is a plasmid DNA encoding Aequorea
CC victoria blue fluorescent protein (BFP) Cycle 3 gene. This sequence is
XX used to illustrate the method of the invention.
SQ Sequence 3637 BP; 942 A; 891 C; 870 G; 934 T; 0 U; 0 Other;
Query Match 72.8%; Score 2665.4; DB 13; Length 3637;
Best Local Similarity 83.7%; Pred. No. 0;
Matches 3065; Conservative 0; Mismatches 571; Indels 27; Gaps 3;
QY 1 CTAAATTGTAAGCGTTAATAATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTC 60
DB |||||
3637 CTAAATTGTAAGCGTTAATAATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTC 3578
QY 61 ATTTTAAACAAATAGGCGGAAATCGGCAAAATCCCTTATAAATCAAAAGAAATAGACCGA 120
DB |||||
3577 ATTTTAAACAAATAGGCGGAAATCGGCAAAATCCCTTATAAATCAAAAGAAATAGACCGA 3518
QY 121 GATAGGGTTGAGTGTGTTCCAGTTTGGAAACAAAGAGTCCACTATTAAAGAAAGCTGGACTC 180
DB |||||
3517 GATAGGGTTGAGTGTGTTCCAGTTTGGAAACAAAGAGTCCACTATTAAAGAAAGCTGGACTC 3458
QY 181 CAAAGTCAAAGGGCGGAAACCCGCTCTATCAGGGCGATGGCCCACTACGTGAACCAATCACC 240
DB |||||
3457 CAAAGTCAAAGGGCGGAAACCCGCTCTATCAGGGCGATGGCCCACTACGTGAACCAATCACC 3398
QY 241 CTAAATCAAGTTTGTGGGTGAGTGCCTTAAGCACTAAATCGGAACCTTAAAGGGAG 300
DB |||||
3397 CTAAATCAAGTTTGTGGGTGAGTGCCTTAAGCACTAAATCGGAACCTTAAAGGGAG 3338
QY 301 CCCCGATTTAGAGCTTCACGGGAAAGCGCGCAAACTGGCGAGAAAGAGGGAAGAA 360
DB |||||
3337 CCCCGATTTAGAGCTTCACGGGAAAGCGCGCAAACTGGCGAGAAAGAGGGAAGAA 3278
QY 361 AGCGAAAGAGCGGCGCTAGGCGCTGGCAAGTGTAGCGGTACGCTGCGGTAAACCAAC 420
DB |||||
3277 AGCGAAAGAGCGGCGCTAGGCGCTGGCAAGTGTAGCGGTACGCTGCGGTAAACCAAC 3218
QY 421 CACACCGCGCGCTTAATGCGCGCTACAGGGCGGTCCTTCCATTCGCATTCAGGCTGGC 480
DB |||||
3217 CACACCGCGCGCTTAATGCGCGCTACAGGGCGGTCCTTCCATTCGCATTCAGGCTGGC 3158
QY 481 CAACTGTTGGGAAGGCGCATCGGTGCGGCGCTTTCGCTATTACGCCAGTGGCGAAAGG 540
DB |||||
3157 CAACTGTTGGGAAGGCGCATCGGTGCGGCGCTTTCGCTATTACGCCAGTGGCGAAAGG 3098
QY 541 GGGATGTGCTCAAGGCGGATTAAGTTGGGTAAACGCCAGGTTTTTCCAGTACAGAGTTG 600
DB |||||
3097 GGGATGTGCTCAAGGCGGATTAAGTTGGGTAAACGCCAGGTTTTTCCAGTACAGAGTTG 3038
QY 601 TAAAAAGAGCGGCGAGTGGCGGCTCGTTCACTTCACTGTTTTTGAACCGGTGAGGAGCGG 660
DB |||||
3037 TAAAAAGAGCGGCGAGTGGCGGCTCGTTCACTTCACTGTTTTTGAACCGGTGAGGAGCGG 2978
QY 661 GCAGACTCGCGTGCAAAATGTGTTTTACAGCGTGTAGGAGCAGATCAAGATCTCTCGACAC 720
DB |||||
2977 CCGCGGTGCGGCGGCTCTAGAACTAGTGATCCCGCGGCTGCGAGGATTTCTTATTGT 2918
QY 721 GCTGCAGAAACAGCAGCTAGATTAAACCTAGAAAGATAATCATATTGTGACGTGTTAA 780

Db 2917 ATAGTTCAATCCATGCCATGTGTAATCCACGACGACGATTCAAACTCAAGAGACCATGT 2858
QY 781 AGATAATC-ATGCGTAAATTTGACGATGGATCTGTAAATACGACTCACTATATAGGCGGAA 839
Db 2857 GGTACGCTTTTGGTGGATCTTTCGAAAGGCGAGATTGTGCGACAGTAATGGTGT 2798
QY 840 TTGGGTACCGGGCCCCCTCGAGGTGCGAGGTATCGATAAGCTTGATATCGAATTCCTG 899
Db 2797 CTGGTAAAGGACAGGGCCATCGCCAAATGGAGTATTTTGTGATAATGGTCTGTAGTT 2738
QY 900 CAGCCCGGGGATCCACTAGTTCTAGAGCGCGCCACCGCGTGGAGTCCAGCTTTTG 959
Db 2737 GAAGGATCCATCTTCAATGTTGTGGGAATTTTGAAGTTAGTTTGTATCCATCTCTTT 2678
QY 960 TTCCCTTTAGTGAAGGTTAATATAGATCCATGCGTCAATTTTACGAGACTATCTTTCTA 1019
Db 2677 GTTTGTCTGCGTGATATACATTTGTGTAGTTAAAGTTGTACTCGAGTTGTGTCCGA 2618
QY 1020 GGGTTAATCTAGCTGATCAGGATCATATCGTGGGTCTTTTTCGGGTCTAGTCAATCGC 1079
Db 2617 GAATGTTTCCATCTTCTTTAAATCAATACCTTTTAACTCGATAC-----GATTAACAAG 2563
QY 1080 CCAAGCTGGCGTATCTGGGCATCGGGAGGAAGCCGCTTTTCCCGCGAGTT 1139
Db 2562 GGTATCACCTTCAACTTGACTTCAGACGCGTCTGTAGTTCCCGTCACTTTTGAAGA 2503
QY 1140 GAAGCGGCATGAAAGAGTTTTCGCGAGGATGACTGCTGCTGCAATTAACCTTTGAGCGAAAA 1199
Db 2502 TATAGTGGTTCCTGTACATACTTCGGGCATGGCACTTTGAAAAGTCATGCCGTTT 2443
QY 1200 CGCACGTTTACATGATGATTCGGGAAGGTGTGGCCATGACGCCCTTTAAACGGTGAAGT 1259
Db 2442 CATATGATCGGATAACGAGAAAGCAITGAACACCATGAGAGAAAGTAGTGACAAAGTGT 2383
QY 1260 TTGGTTACGCCACCTGGGATACCACTGCTGCGCGCTTTTCGGACACAGTTCCGGATG 1319
Db 2382 TGGCCATGGAACHAGTAGTTTCCAGTAGTGCAATAAATTTAAGTGTAAAGCTTTCCGTA 2323
QY 1320 GTACGCGGAAGCGCATCAGCAACCCGAAACATACCGGGACAGCGGAACTGCCGTGCC 1379
Db 2322 GTAGCATCACCTTCACCTCTCCACTGACAGAAATTTGTGCCATTAACATCACCATC 2263
QY 1380 GGTGTGCAGATTAATGACAGCGGTGCGCGCTGGGATATTACGTACGAGAGACGGGTAT 1439
Db 2262 TAATTCACAAGAAT-----TGGACAACTCCAGTGAAAGATTTC 2224
QY 1440 CTTGGCTGGATGCGCAGAAATGACATGGATACCCGTTAGTACCCCGCGCGCGCT 1499
Db 2223 TTCCTCTTACTCATCGGTACCCAGCTTTTGTTCCTTTAGTGAGGGTTAATTTGCCGCT 2164
QY 1500 TGGCGTAAATCATGGTCATAGCTGTTTCTGTGTGAAATTTTATCCGCTCACAATTTCCAC 1559
Db 2163 TGGCGTAAATCATGGTCATAGCTGTTTCTGTGTGAAATTTTATCCGCTCACAATTTCCAC 2104
QY 1560 ACAACATACGAGCCGGAAGATAAAGTGTAAAGCCTGGGGTGCTTAATGAGTGAGCTAAC 1619
Db 2103 ACAACATACGAGCCGGAAGATAAAGTGTAAAGCCTGGGGTGCTTAATGAGTGAGCTAAC 2044
QY 1620 TCACATTAATGGTTGGCTCACTGCCGCTTTCCAGTCCGGAACCTGTGTCGCGAG 1679
Db 2043 TCACATTAATGGTTGGCTCACTGCCGCTTTTCCAGTCCGGAACCTGTGTCGCGAG 1984
QY 1680 TGCATTAATGAATTCGGCCAAACGCGGGGAGAGCGGTTTGTGCTATTTGGCGCTCTTTCCG 1739
Db 1983 TGCATTAATGAATTCGGCCAAACGCGGGGAGAGCGGTTTGTGCTATTTGGCGCTCTTTCCG 1924
QY 1740 CTTTCTCGCTCACTGACTCGTGGCTCGGTTCGGTTCGGTGGCGGAGCGGTATCAGCTC 1799
Db 1923 CTTTCTCGCTCACTGACTCGTGGCTCGGTTCGGTTCGGTGGCGGAGCGGTATCAGCTC 1864
QY 1800 ACTCAAGGGGGTAAATACGGTTATCCACAGAAATCAGGGATTAACGAGAAAGACATGT 1859

Db 1863 ACTCAAGGGGGTAAATACGGTTATCCACAGAAATCAGGGATTAACGAGAAAGAACATGT 1804
QY 1860 GAGCAAAAAGCCAGCAAAAAGCCAGAACCGTAAAAAAGCCGGTGTGTCGGCTTTTTC 1919
Db 1803 GAGCAAAAAGCCAGCAAAAAGCCAGAACCGTAAAAAAGCCGGTGTGTCGGCTTTTTC 1744
QY 1920 ATAGGCTCGCCCTCGACGAGCATCAAAAATCGACGCTCAAGTCAAGAGTGGCGAA 1979
Db 1743 ATAGGCTCGCCCTCGACGAGCATCAAAAATCGACGCTCAAGTCAAGAGTGGCGAA 1684
QY 1980 ACCGACAGGACTATAAAGATACAGGCGTTTCCCTCGAAAGCTCCCTCGTCGCTCTC 2039
Db 1683 ACCGACAGGACTATAAAGATACAGGCGTTTCCCTCGAAAGCTCCCTCGTCGCTCTC 1624
QY 2040 CTGTTCCGACCTCGCGCTTACCGGATACCTGTCCGCCCTTTCTCTCGGGAAGCGTGG 2099
Db 1623 CTGTTCCGACCTCGCGCTTACCGGATACCTGTCCGCCCTTTCTCTCGGGAAGCGTGG 1564
QY 2100 CGCTTCTCATAGCTACAGCTGTAGGTATCTCAGTTCCGCTGAGGTGCTTCGCTCCAAGC 2159
Db 1563 CGCTTCTCATAGCTACAGCTGTAGGTATCTCAGTTCCGCTGAGGTGCTTCGCTCCAAGC 1504
QY 2160 TGGGCTGTGTGACGAAACCCCGCTTCAGCCCGAACCGCTCGCGCTTTATCCGGTAACTATC 2219
Db 1503 TGGGCTGTGTGACGAAACCCCGCTTCAGCCCGAACCGCTCGCGCTTTATCCGGTAACTATC 1444
QY 2220 GTCTTGAFTCCAAACCGGTAAGACAGACTTATTCGCACTGCGAGCAGCACCTGGTAAACA 2279
Db 1443 GTCTTGAFTCCAAACCGGTAAGACAGACTTATTCGCACTGCGAGCAGCACCTGGTAAACA 1384
QY 2280 GATTTAGCAGAGGGATGTAGGGGTGCTACAGAGTTCTTGAGTGGTGGCTTAACCT 2339
Db 1383 GATTTAGCAGAGGGATGTAGGGGTGCTACAGAGTTCTTGAGTGGTGGCTTAACCT 1324
QY 2340 ACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCAGCTTACCTTCG 2399
Db 1323 ACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCAGCTTACCTTCG 1264
QY 2400 GAAAAAGAGTGGTAGCTCTTGATCCGCAAAACAAACCAACCCGCTGGTAGCGGTGTTTT 2459
Db 1263 GAAAAAGAGTGGTAGCTCTTGATCCGCAAAACAAACCAACCCGCTGGTAGCGGTGTTTT 1204
QY 2460 TTGTTTGAAGCAGCAGATTAACCGCAGAAAGGATCTCAAGAGATCTTTCATCT 2519
Db 1203 TTGTTTGAAGCAGCAGATTAACCGCAGAAAGGATCTCAAGAGATCTTTCATCT 1144
QY 2520 TTTCTACGGGCTGCAAGCTCAGTGGAAACGAAACCTCAGCTTAAAGGATTTTGGTCACTGA 2579
Db 1143 TTTCTACGGGCTGCAAGCTCAGTGGAAACGAAACCTCAGCTTAAAGGATTTTGGTCACTGA 1084
QY 2580 GATTTCAAAAAGGATCTTACCTAGATCTTTTAAATTTAAATTTAAATTTAAATTTAAATTTAA 2639
Db 1083 GATTTCAAAAAGGATCTTACCTAGATCTTTTAAATTTAAATTTAAATTTAAATTTAAATTTAA 1024
QY 2640 TCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTTACCAATGCTTAAATCAGTGAGGCAC 2699
Db 1023 TCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTTACCAATGCTTAAATCAGTGAGGCAC 964
QY 2700 CTATCTCAGGATCTGTCTATTTTCTCATCTAGTTTGCCTGACTCCCGCTCGTGTAGA 2759
Db 963 CTATCTCAGGATCTGTCTATTTTCTCATCTAGTTTGCCTGACTCCCGCTCGTGTAGA 904
QY 2760 TAACTACGATACGGGAGGCTTACCATCTGGCCCCCAGTCTCGCAATGATACCGCGAGACC 2819
Db 903 TAACTACGATACGGGAGGCTTACCATCTGGCCCCCAGTCTCGCAATGATACCGCGAGACC 844
QY 2820 CAGCTCACCGGCTCCAGATTTTATCAGCAATAAACCCAGCAGCCCGGAAAGGCGGAGCGCA 2879
Db 843 CAGCTCACCGGCTCCAGATTTTATCAGCAATAAACCCAGCAGCCCGGAAAGGCGGAGCGCA 784
QY 2880 GAAGTGGTCTCCAACTTTTATCCGCTCCATCCAGTCTATTAATTTCTCCGGGAAGCTA 2939
Db 783 GAAGTGGTCTCCAACTTTTATCCGCTCCATCCAGTCTATTAATTTCTCCGGGAAGCTA 724

PF	01-FEB-2002; 2002WO-US003055.		
XX			
PR	02-FEB-2001; 2001US-0266386P.		
PR	14-FEB-2001; 2001US-0268785P.		
XX			
PA	(LARG-) LARGE SCALE BIOLOGY CORP.		
XX			
PI	Padgett HS, Fitzmaurice WP, Lindo JA;		
XX			
DR	WPI; 2003-092898/08.		
XX			
PT	Increasing complementarity by making sequence variants from heteroduplex		
PT	polynucleotides, useful for producing improved gene products from		
PT	randomly mutated genes.		
PS	Example 5; Page 99-100; 105pp; English.		
XX			
CC	The invention relates to an in vitro method of making sequence variants		
CC	from at least one heteroduplex polynucleotide that has at least two-non		
CC	complementary nucleotide base pairs. The method comprises combining a		
CC	heteroduplex polynucleotide with CEL I, T4 DNA polymerase and T4 DNA		
CC	ligase, or an agent or agents with exonuclease activity and allowing		
CC	sufficient time for the percentage of complementarity to increase, where		
CC	one or more variants are made. The method can also be used to increase		
CC	diversity in a population of sequences. The method is useful for		
CC	producing improved gene products from randomly mutated genes or from		
CC	expression from a suitable plant, animal, fungal, yeast or bacterial		
CC	expression vector, and provides a high-efficiency recovery of		
CC	recombinants for generating novel polynucleotides from parental templates		
CC	that are more diverse and with a lower percentage of sequence identity.		
CC	This sequence represents plasmid DNA, used in heteroduplex substrate		
CC	preparation		
XX			
SQ	Sequence 3637 BP; 942 A; 892 C; 870 G; 933 T; 0 U; 0 Other;		
Query Match 72.8%; Score 2664.8; DB 8; Length 3637;			
Best Local Similarity 83.8%; Pred. No. 0;			
Matches 3070; Conservative 0; Mismatches 567; Indels 25; Gaps 4;			
QY	1	CTAATTTGTAAGCGTTAATATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTC	60
Db	3637	CTAAATTTGTAAGCGTTAATATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTC	3578
QY	61	ATTTTTTAAACCAATAGCGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGATAGACCGA	120
Db	3577	ATTTTTTAAACCAATAGCGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGATAGACCGA	3518
QY	121	GATAGGGTTGAGTGTGTTCCAGTTTGGAAACAAGTCCACTATTTAAAGAACGTGGACTC	180
Db	3517	GATAGGGTTGAGTGTGTTCCAGTTTGGAAACAAGTCCACTATTTAAAGAACGTGGACTC	3458
QY	181	CAACGTCAAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC	240
Db	3457	CAACGTCAAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC	3398
QY	241	CTAATCAAGTTTTTTTGGGGTCCAGGTCCCGTAAGCACTAAATCGGAACCCCTAAAGGGAG	300
Db	3397	CTAATCAAGTTTTTTTGGGGTCCAGGTCCCGTAAGCACTAAATCGGAACCCCTAAAGGGAG	3338
QY	301	CCCCCGATTAGAGCTTGACGGGGAAAGCCGCGCAACGTGGCGAGAAAGGAAGGAAGAA	360
Db	3337	CCCCCGATTAGAGCTTGACGGGGAAAGCCGCGCAACGTGGCGAGAAAGGAAGGAAGAA	3278
QY	361	AGCGAAAGAGCGGCGCTAGGCGCTGCGCAAGTCTAGCGGTACCGCTGCGCTTAACCCAC	420
Db	3277	AGCGAAAGAGCGGCGCTAGGCGCTGCGCAAGTCTAGCGGTACCGCTGCGCTTAACCCAC	3218
QY	421	CACACCCCGCGCTTAATCGCCGCTACAGGGCGCGTCCCATTCGCCATTTCAGGCTGCG	480
Db	3217	CACACCCCGCGCTTAATCGCCGCTACAGGGCGCGTCCCATTCGCCATTTCAGGCTGCG	3158
QY	481	CAACTGTTGGGAAGGGCGATCGGTGGGGCTCTTCGCTATTATACGCCAGCTGGCGAAGG	540

Db 3157 CAACTGTTGGGAAGGGGATCGGTGCGGGCTCTTCGCTATTACGCCAGCTGCGGAAAGG 3098
Qy 541 GGGATGTGTCGAAGCGGATTAAGTTGGGTAAAGCGCAGGGTTTCCAGTCAAGAGCTTG 600
Db 3097 GGGATGTGTCGAAGCGGATTAAGTTGGGTAAAGCGCAGGGTTTCCAGTCAAGAGCTTG 3038
Qy 601 TAAACGAGCGGCAGTGAAGGGCTCGTTCAATCAAGTTTGAACCCGTGGAGACGG 660
Db 3037 TAAACGAGCGGCAGTGAAGGGCTCGTTCAATCAAGTTTGAACCCGTGGAGTCCA 2978
Qy 661 GCAGACTCGCGGTGCAAAATGTGTTTACAGCGTGAATGGAGCAGATGAAGATGCTCCAGAC 720
Db 2977 CCGCGTGGCGG-----CGCTCTAGAACTAGTGTGATCCCC--CGG 2939
Qy 721 GTGCGAGAACCGACGCTAGATTAAACCTAGAAAGATAATCATATTGTGACGTACGTTAA 780
Db 2938 GTGCGAGGAATTCATTATTTGTATAGTTCAATGCCATGTGTAATCCCGCAGCAGATT 2879
Qy 781 AGATAATCATGCTAAATTTGACGATGGATCTGTAATACGACTCACTATAGGGCGAAT 840
Db 2878 CAAACTCAAGAGGACCATGTGGTCAAGCTTTTCGTTGGGATCTTTGGAAGGGCAGATT 2819
Qy 841 TGGGTACCGGGCCCCCTCGAGTTCGACGGTATCGATAAGCTTGTATATCGAATTCCTGC 900
Db 2818 GTGTCAGAGTAATGGTTGTCTGGTTAAAGACAGGGCCATCGCCAAATTGAGTATTTT 2759
Qy 901 AGCCCGGGGATCCACTAGTTCTAGAGCGCGCCACCGCGGTGGAGTCCAGCTTTTGT 960
Db 2758 GTTGATAATGGTCTGTAGTTGAACGATCCATCTTCAATGTTGTGGCGAAATTTTGAAGT 2699
Qy 961 TCCCTTTAGTAGGGTTAATTAGATCCCATGGTCAATTTTACGGCAGACTATCTTCTAG 1020
Db 2698 TAGCTTTGATTCATTTCTTTGTTGTTCTGCGGTGATGTATACATTTGTGTAG--TTATAG 2640
Qy 1021 GGTAAATCTAGCTGCATCAGGATCATATCGTGGGTCTTTTTCGGGCTCAGTCACGCC 1080
Db 2639 TTGTACTCAGTTTGTGTCGGAGATGTTTTCATCTCTTTAAATCAATACCTTTTAAAC 2580
Qy 1081 CAAAGTGGCGCTATCTGGGCATCGGGAGGAGAAAGCCGTGCTTTTCCCGCAGGTTG 1140
Db 2579 TCGATACGATTAAACAGGGTATCACCTTCAAACTTGACTTCAGCAGCGCTTTGTAGTT 2520
Qy 1141 AAGCGCATGGAAGATTGTCGGAGGATGACTGCTGCTGCAATGAGTTGAGCGGAAC 1200
Db 2519 CCGTCACTTTTGAAGATATAGTGCCTTCTGTACATAACCTTCGGGCATGGCACTCTTG 2460
Qy 1201 GCACGTTTACCATGATTTCGGGAGGTGTGCGCATGCACGCTTTTACGGTGAACCT 1260
Db 2459 AAAAGTCAATCGCGTTTCATATGATCCGGATTAACGGGAAAGCAATGAACACCAATAGAG 2400
Qy 1261 TCGTTTCAGGCCACCTGGGATACAGGTTCTGCGGGCTTTTTCGGGACACAGTTCCGGATGG 1320
Db 2399 AAAGTAGTGACAAGTTGCGCCATGGAACAGGTAGTTTCCAGTAGTGCAATAAATTTA 2340
Qy 1321 TCAGCCGGAAGCGCATCAGCAACCCGAAACAATACGGCGACAGCGGAACTCGCGTGGCG 1380
Db 2339 AGGGTAAGCTTTCCGTATGTAGCATCACCTTCACCTCTCCACTGACGAAAAAATTTGTC 2280
Qy 1381 GTGTGCAGATTAAAGACAGGGTGGCGCTGGGATTTACGTCAAGCAGGAGCGGGTATC 1440
Db 2279 CCATTAAACATACCATTAAATTCACAAAGAAATGGGACAACTCCAGTGAAGATTTCTT 2220
Qy 1441 CTGGCTGGATGCGCGAAGAAATGACATGATATCCCGTGAGTTTACCCGCGGGCGCGCTT 1500
Db 2219 C---CTTTACTCATCGGTACCCAGCTTTTGTTCCTTTTAGTGAGGGTTAATTGCGGCTT 2163
Qy 1501 GCGTAATCATGTGTATAGCTGTTTCTGTGTGAATTTGTATTCGCTCACAAATTCACA 1560
Db 2162 GCGTAATCATGTGTATAGCTGTTTCTGTGTGAATTTGTATTCGCTCACAAATTCACA 2103
Qy 1561 CAACTACAGCGGAGCATAAAGTAAAGCTGGGTGCTTAATGATGAGCTTAAT 1620
Db 2102 CAACTACAGCGGAGCATAAAGTAAAGCTGGGTGCTTAATGATGAGCTTAAT 2043

Qy 1621 CACATTAAATTTGGTTGCGCTCACTGCCCCGCTTTTCAAGTCGGGAAACCTGTGTGTCAGCT 1680
Db 2042 CACATTAAATTTGGTTGCGCTCACTGCCCCGCTTTTCAAGTCGGGAAACCTGTGTGTCAGCT 1983
Qy 1681 GCATTAAATCAATCGCCAAACGCGCGGAGAGCGGTTTGGTATTGCGGCTCTTCCGC 1740
Db 1982 GCATTAAATCAATCGCCAAACGCGCGGAGAGCGGTTTGGTATTGCGGCTCTTCCGC 1923
Qy 1741 TTCTCGCTCACTGACTCGCTCGCTCGGTTCGGTTCGGCGAGCGGTATCAGCTCA 1800
Db 1922 TTCTCGCTCACTGACTCGCTCGCTCGGTTCGGTTCGGCGAGCGGTATCAGCTCA 1863
Qy 1801 CTCAAAGGGGTAAATACGGTTATCCACAGAAATCAGGGATTAACGCAAGGAAAGAAATGTG 1860
Db 1862 CTCAAAGGGGTAAATACGGTTATCCACAGAAATCAGGGATTAACGCAAGGAAAGAAATGTG 1803
Qy 1861 AGCAAAAGCCAGCAAAAGCCAGGAACCGTAAAGAGCCGGTTCCTGCGTTTTTCCA 1920
Db 1802 AGCAAAAGCCAGCAAAAGCCAGGAACCGTAAAGAGCCGGTTCCTGCGTTTTTCCA 1743
Qy 1921 TAGGCTCCGCCCTCTGACGAGCATCAAAAATTCAGCTCAAGTCAGAGGTGGGAAA 1980
Db 1742 TAGGCTCCGCCCTCTGACGAGCATCAAAAATTCAGCTCAAGTCAGAGGTGGGAAA 1683
Qy 1981 CCGCAGAGCATATAAAGATAACAGCGTTTCCCCCTGGAGCTCCTCGTGGCTCTCC 2040
Db 1682 CCGCAGAGCATATAAAGATAACAGCGTTTCCCCCTGGAGCTCCTCGTGGCTCTCC 1623
Qy 2041 TGTTCGAGCCCTGCGCTTACCGGATACCTGTCGGCTTTCTCCCTCGGAGCGTGGC 2100
Db 1622 TGTTCGAGCCCTGCGCTTACCGGATACCTGTCGGCTTTCTCCCTCGGAGCGTGGC 1563
Qy 2101 GCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTTCGGTGTAGGTCTGCTCGCTCAAGCT 2160
Db 1562 GCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTTCGGTGTAGGTCTGCTCGCTCAAGCT 1503
Qy 2161 GGGCTGTGTGACGAAACCCCTGTCAGCCGACCGCTCGGCTTTATCCGGTAACTATCG 2220
Db 1502 GGGCTGTGTGACGAAACCCCTGTCAGCCGACCGCTCGGCTTTATCCGGTAACTATCG 1443
Qy 2221 TCTTGAGTCCACCCGTTAAGACAGCATTTATGCCACTGCGCAGCAGCCACTGTTAAACAG 2280
Db 1442 TCTTGAGTCCACCCGTTAAGACAGCATTTATGCCACTGCGCAGCAGCCACTGTTAAACAG 1383
Qy 2281 GATTAGCAGAGCAGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTAACTA 2340
Db 1382 GATTAGCAGAGCAGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTAACTA 1323
Qy 2341 CCGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTTACCTCGG 2400
Db 1322 CCGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTTACCTCGG 1263
Qy 2401 AAAAGAGTTGTAGTCTTGTATCCGGGAAACAACACCGCTGGTAGCGGTGTTTTT 2460
Db 1262 AAAAGAGTTGTAGTCTTGTATCCGGGAAACAACACCGCTGGTAGCGGTGTTTTT 1203
Qy 2461 TGTGTCAGCAGCAGATTACCGCAGAAAAAAGGATCTCAAGAAAGATCTTTGATCTT 2520
Db 1202 TGTGTCAGCAGCAGATTACCGCAGAAAAAAGGATCTCAAGAAAGATCTTTGATCTT 1143
Qy 2521 TTCTACGGGGTCTGACGCTCAGTGGAAAGAAACCTCACTTAAGGGATTTTGGTCAATGAG 2580
Db 1142 TTCTACGGGGTCTGACGCTCAGTGGAAAGAAACCTCACTTAAGGGATTTTGGTCAATGAG 1083
Qy 2581 ATTATCAAAAAGGATCTTCACCTAGATCTTTTAAATTTAAATTTAAATCAAT 2640
Db 1082 ATTATCAAAAAGGATCTTCACCTAGATCTTTTAAATTTAAATTTAAATCAAT 1023
Qy 2641 CTAAAGTATATAGTAAACCTTGGTCTGACAGTTTACCAATCTTAATCAGTGAGGCACC 2700
Db 1022 CTAAAGTATATAGTAAACCTTGGTCTGACAGTTTACCAATCTTAATCAGTGAGGCACC 963

QY	2701	TATCTCAGCGATCTGTCTATTTTCGTTCACTCCATAGTGTGCTGACTCCCGTCTGTGTAGAT	2760
Db	962	TATCTCAGCGATCTGTCTATTTTCGTTCACTCCATAGTGTGCTGACTCCCGTCTGTGTAGAT	903
QY	2761	AACTAGGATACGGAGGGCTTACCATCTGCGCCCGAGTGTGCAATGATACCCGAGACCC	2820
Db	902	AACTAGGATACGGAGGGCTTACCATCTGCGCCCGAGTGTGCAATGATACCCGAGACCC	843
QY	2821	AGCTCACCGGCTCCAGATTTATCAGCAATAAACAGCAGCGGAGGCGCGAGCGAG	2880
Db	842	AGCTCACCGGCTCCAGATTTATCAGCAATAAACAGCAGCGGAGGCGCGAGCGAG	783
QY	2881	AAGTGGTCTCGCACTTTATCCGCTCCATCCAGTCTATTAATTTGTCGGGAAAGTAG	2940
Db	782	AAGTGGTCTCGCACTTTATCCGCTCCATCCAGTCTATTAATTTGTCGGGAAAGTAG	723
QY	2941	AGTAAGTAGTTCGCCAGTTAATAGTTTGGCGCAACGTTGTTGCCATTTGCTACAGGCATCGT	3000
Db	722	AGTAAGTAGTTCGCCAGTTAATAGTTTGGCGCAACGTTGTTGCCATTTGCTACAGGCATCGT	663
QY	3001	GGTGTACGCTCGTCTGTTGGTATGGCTTCACTCAGCTCCGTTCCCAACGATCAAGGCG	3060
Db	662	GGTGTACGCTCGTCTGTTGGTATGGCTTCACTCAGCTCCGTTCCCAACGATCAAGGCG	603
QY	3061	AGTTACATGATCCCCATGTTGTGCAAAAAGCGGTTAGCTCTTCGTTCTCCGATCGT	3120
Db	602	AGTTACATGATCCCCATGTTGTGCAAAAAGCGGTTAGCTCTTCGTTCTCCGATCGT	543
QY	3121	TGTCAGAGTAAGTTGGCGCGAGTGTATCACTCATGTTTATGGCAGCAGCTGCATAATTC	3180
Db	542	TGTCAGAGTAAGTTGGCGCGAGTGTATCACTCATGTTTATGGCAGCAGCTGCATAATTC	483
QY	3181	TCCTTACTGTCAGCCATCGTAAAGTGTCTTTCTGACTGTGTAGTCACTCAACCAAGTC	3240
Db	482	TCCTTACTGTCAGCCATCGTAAAGTGTCTTTCTGACTGTGTAGTCACTCAACCAAGTC	423
QY	3241	ATTTGAGAATAGTGTATGCGCGACCGAGTTGCTCTTGGCGGGGTCAATAACGGGATAA	3300
Db	422	ATTTGAGAATAGTGTATGCGCGACCGAGTTGCTCTTGGCGGGGTCAATAACGGGATAA	363
QY	3301	TAACGGCGCACATAGCAGAACCTTTAAAGTGCTCATTTGAAAAACGTTCTTCGGGGG	3360
Db	362	TAACGGCGCACATAGCAGAACCTTTAAAGTGCTCATTTGAAAAACGTTCTTCGGGGG	303
QY	3361	AAAACTCTCAAGATCTTACCGCTGTGAGATCCAGTTTCGATTAACCCACCTCGTGACC	3420
Db	302	AAAACTCTCAAGATCTTACCGCTGTGAGATCCAGTTTCGATTAACCCACCTCGTGACC	243
QY	3421	CAACTGATCTTCAGCATCTTTTACTTTTCAACAGCGTTTCTGGGTGAGCAAAAACAGGAAG	3480
Db	242	CAACTGATCTTCAGCATCTTTTACTTTTCAACAGCGTTTCTGGGTGAGCAAAAACAGGAAG	183
QY	3481	GCAAAATGCGCAAAAAGGGAATAGGCGGACAGGAAATGTTCAATACTCATACTCTT	3540
Db	182	GCAAAATGCGCAAAAAGGGAATAGGCGGACAGGAAATGTTCAATACTCATACTCTT	123
QY	3541	CTTTTTCATATTAATTAAGACATTTATTCAGGTTTATTTGCTCATGAGCGGATACATAT	3600
Db	122	CTTTTTCATATTAATTAAGACATTTATTCAGGTTTATTTGCTCATGAGCGGATACATAT	63
QY	3601	TGAATGTATTTAGAAAAATAAACAATAGGGGTTCCGGGCAATTTCCCGCAAAAGTGCC	3660
Db	62	TGAATGTATTTAGAAAAATAAACAATAGGGGTTCCGGGCAATTTCCCGCAAAAGTGCC	3
QY	3661	AC 3662	
Db	2	AC 1	

RESULT 7
ADM68448/c
ID ADM68448 standard; DNA; 3637 BP.
XX

AC	ADM68448;	
XX	03-JUN-2004 (first entry)	
DE	Cycle 3 green fluorescent protein, GFP, DNA construct.	
XX	ds; mismatch endonuclease; endonuclease; gene shuffling technology;	
KW	single nucleotide polymorphism; cancer susceptibility;	
KW	sequence variation redistribution; cycle 3 GFP;	
XX	green fluorescent protein.	
OS	Aequorea victoria.	
XX	Synthetic.	
PN	US2003157682-A1.	
XX	21-AUG-2003.	
PD	31-JAN-2003; 2003US-00356708.	
PF	01-FEB-2002; 2002US-0353722P.	
XX	14-MAR-2002; 2002US-00098155.	
PR	01-AUG-2002; 2002US-00211079.	
XX	(PADG/) PADGETT H S.	
PA	(VAEW/) VAEWHONGS A A.	
PA	(VOJD/) VOJDANI F S.	
PA	(SMIT/) SMITH M L.	
PA	(LIND/) LINDBO J A.	
XX	(FITZ/) FITZMAURICE W P.	
PI	Padgett HS, Vaewhongs AA, Vojdani FS, Smith ML, Lindbo JA;	
PI	Fitzmaurice WP;	
XX	WPI; 2003-766176/72.	
DR	Making a mismatch endonuclease, useful in gene shuffling and in detection	
XX	of single nucleotide polymorphisms, comprises transfecting a host with a	
PT	recombinant viral vector including a polynucleotide encoding a mismatch	
PT	endonuclease.	
XX	Example 9; SEQ ID NO 17; 79pp; English.	
PS	The invention relates to a method of making a mismatch endonuclease	
XX	enzyme comprising transfecting a host plant, animal, yeast, fungus or	
CC	bacterium with a recombinant viral vector that encodes a polynucleotide	
CC	sequence for a mismatch endonuclease, growing the host so that the	
CC	polynucleotide is expressed, and extracting the mismatch endonuclease	
CC	enzyme from the host. The method is useful for making mismatch	
CC	endonuclease enzymes, for obtaining peptides and polynucleotides with	
CC	desired functional properties and for detecting mutations. The mismatch	
CC	endonuclease enzymes are useful in gene shuffling technology for	
CC	developing new genes, in detecting single nucleotide polymorphisms for	
CC	e.g. detecting evidence of cancer susceptibility, or in redistributing	
CC	sequence variations between non-identical polynucleotide sequences. The	
CC	present sequence represents the cycle 3 green fluorescent protein, GFP,	
CC	DNA construct.	
XX	Sequence 3637 BP; 942 A; 892 C; 870 G; 933 T; 0 U; 0 Other;	
SQ	Query Match 72.8%; Score 2664.8; DB 11; Length 3637;	
	Best Local Similarity 83.8%; Pred. No. 0;	
	Matches 3070; Conservative 0; Mismatches 567; Indels 25; Gaps 4;	
QY	1 CTAAATTCGATCGTTAATATTTTGTAAATTCGCGTAAATTTTGTAAATCAGCTC 60	
Db	3637 CTAAATTCGATCGTTAATATTTTGTAAATTCGCGTAAATTTTGTAAATCAGCTC 3578	
QY	61 ATTTTAAACCAATAGGCGGAAATCGGCAAAATCCCTTATATAATCAAAGATAGACCGA 120	
Db	3577 ATTTTAAACCAATAGGCGGAAATCGGCAAAATCCCTTATATAATCAAAGATAGACCGA 3518	
QY	121 GATAGGGTTGAGTGTGTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGACGTGGACTC 180	

CC of improved polynucleotide sequences possessing at least one desired
CC phenotypic characteristic (e.g., promotes transcription of linked
CC polynucleotides), where such polynucleotides are useful for expression
CC from a plant, animal, fungal, yeast, or bacterial expression vector, for
CC integration to form a transgenic plant, animal or microorganism, and for
CC expression of a ribozyme. This sequence represents DNA used in the scope
XX of the invention.

SQ Sequence 3637 BP; 942 A; 892 C; 870 G; 933 T; 0 U; 0 Other;

```
Query Match          72.8%; Score 2664.8; DB 12; Length 3637;
Best Local Similarity 83.8%; Pred. No. 0;
Matches 3070; Conservative 0; Mismatches 567; Indels 25; Gaps 4;

QY 1 CTAATTGTAAGCGTTAAATTTTGTAAATTCGCGTTAAATTTTGTAAATTCAGCTC 60
DB 3637 CTAATTGTAAGCGTTAAATTTTGTAAATTCGCGTTAAATTTTGTAAATTCAGCTC 3578
QY 61 ATTTTAAACCAATAGCGCGAATCGGCAAAATCCCTTAAATCAAAAGATAGACCGA 120
DB 3577 ATTTTAAACCAATAGCGCGAATCGGCAAAATCCCTTAAATCAAAAGATAGACCGA 3518
QY 121 GATAGGTTGAGTGTGTTTCCAGTTTGGAAACAGAGTCCACTATTAAAGAACTGGACTC 180
DB 3517 GATAGGTTGAGTGTGTTTCCAGTTTGGAAACAGAGTCCACTATTAAAGAACTGGACTC 3458
QY 181 CAAAGCGTCAAGGCGGCAAAACCGTCTATCAGGCGATGGCCCACTACGTGAACCAATCACC 240
DB 3457 CAAAGCGTCAAGGCGGCAAAACCGTCTATCAGGCGATGGCCCACTACGTGAACCAATCACC 3398
QY 241 CTAATCAAGTTTTTGGGTCGAGGTGCGGTAAAGCACTAAATCGGAACCCCTAAAGGGAG 300
DB 3397 CTAATCAAGTTTTTGGGTCGAGGTGCGGTAAAGCACTAAATCGGAACCCCTAAAGGGAG 3338
QY 301 CCCCAGATTTAGAGCTTTCGCGGGAAGCCGCGAAGCTGGCGAGAAAGGAAGGA 360
DB 3337 CCCCAGATTTAGAGCTTTCGCGGGAAGCCGCGAAGCTGGCGAGAAAGGAAGGA 3278
QY 361 AGCGAAGGAGGCGGCGCTAGGCGCTGGCAAGTGTAGCGGTACAGCTGCGGTAAACAC 420
DB 3277 AGCGAAGGAGGCGGCGCTAGGCGCTGGCAAGTGTAGCGGTACAGCTGCGGTAAACAC 3218
QY 421 CACACCCGCGGCTTAATGCGCGCTACAGGCGCGTCCCATTCGCCATTCAGGCTGCG 480
DB 3217 CACACCCGCGGCTTAATGCGCGCTACAGGCGCGTCCCATTCGCCATTCAGGCTGCG 3158
QY 481 CAACTGTTGGGAAGGCGATCGGTGCGGCTCTTCGCTATTACGCCAGCTGCGGAAGG 540
DB 3157 CAACTGTTGGGAAGGCGATCGGTGCGGCTCTTCGCTATTACGCCAGCTGCGGAAGG 3098
QY 541 GGGATGTGTCGAAGCGGATTAAGTTGGGTAAAGCCAGCGGTTTCCAGTCAAGCTGTTG 600
DB 3097 GGGATGTGTCGAAGCGGATTAAGTTGGGTAAAGCCAGCGGTTTCCAGTCAAGCTGTTG 3038
QY 601 TAAACGACGCGGCGAGTGGCGCTCTGTTCAATTCAGTTTGTAAACCCGTGGAGACGG 660
DB 3037 TAAACGACGCGGCGAGTGGCGCTCTGTTCAATTCAGTTTGTAAACCCGTGGAGACGG 2978
QY 661 GCAGATCCGCGTGCAATGTGTTTACAGCGGTGAGGAGCAGATGAAGATGCTGCACAC 720
DB 2977 CCGCGTGGCGG-----CCGCTCTAAGACTAGTGATCCCC--CGG 2939
QY 721 GCTGCAGAACACGCGAGCTAGATTAAACCTAGAAAGATAATCATATTGTGACGTACGTTAA 780
DB 2938 GCTGCAGGAATCTTATTGTTATGTTTTCATCCATGCCATGTGTATCCCGAGCAGTTA 2879
QY 781 AGATAATCATCGGTAAATTTGACGCATGGGATCTGTAAATACGACTCACTATAGGGCGAAT 840
DB 2878 CAAACTCAAGAAAGGACCATGTGGTCACGCTTTTCGTTGGGATCTTTCGAAAGGGCAGATT 2819
QY 841 TGGGTACCGGGCCCCCTCGAGTGCAGCGGTATCGATAAGCTTGATATCGAATTCCTGC 900
DB 2818 GTGTCACAGGTAATGGTTGTCTGGTAAAGGACAGGCGCATCGCCAAATTGGAGTATTTT 2759
```

```
QY 901 AGCCCGGGGATCCACTAGTTCTAGAGCGGCGGCCCAACCGCGTGGAGTCCAGCTTTTGT 960
DB 2758 GTTGATAATGCTGTCTAGTTGAAGCGATCCATCTTCAATGTTGTGGCGAATTTTGAAGT 2699
QY 961 TCCCTTTAGTGAGGGTTAATTAGATCCCATGGTCAATTTTACGCAGACTATCTTCTAG 1020
DB 2698 TAGCTTTGATTCATTTCTTTTGTCTGCGGTGATGTATACATGTTGTGAG-TTATAG 2640
QY 1021 GGTAAATCTAGCTGCATCAGGATCATATCGTGGGCTCTTTTTCGGCTCAGTCATCGCC 1080
DB 2639 TTGTACTCGAGTTTGTGTCGAGAAATGTTTCATCTTCTTTAAATCAATACCTTTTAA 2580
QY 1081 CAAAGTGGCGCTATCTGGGCATCGGGAGGAAGAACCGCGTCTTTTCCCGGAGGTTG 1140
DB 2579 TCGATACGATTAAACAAGGGTATCACCTTCAAACTTGACTTCAGCACGCGCTTGTAGTTC 2520
QY 1141 AAGCGGCATGGAAGAGTTTTCGGGAGATGACTGCTGCTGCAATTCAGTTTGAGCGAAAC 1200
DB 2519 CCGTCATCTTTGAAAGATATAGTGGTTCCTGTACATAAACCTTCGGGCATGGACCTCTTG 2460
QY 1201 GCACGTTTACCATGATGATTCGGGAAGGTGTGGCCATGCAACGCTTTTAAACGGTGAAC 1260
DB 2459 AAAAGTTCATCGCTTTCATATGATCCGATTAACGGGAAGCATTTGAACACCATAAGAG 2400
QY 1261 TCGTTACGCCACCTGGGATACAGTTCTGTCGGGCTTTTCCGGAACACAGTTCCGGATGG 1320
DB 2399 AAAAGTAGTACAAAGTGTGGCCATGGAACAGGTAGTTTTCAGTAGTGCATAAATAA 2340
QY 1321 TCAGCCCGAAGCGCATCAGAACCCGACAAATACCGGCGACAGCCGGAACCTGCCGTCGG 1380
DB 2339 AGGGTAAGCTTTCCGTTATGTAGCATACCTTCAACCTCTCCACTGACAGAAATTTGTGC 2280
QY 1381 GTGTGCAGATTAAATGACAGCGGTGGCGCTGGGATATTACGTGACGAGGAGCGGTATC 1440
DB 2279 CCAITTAACATCACCATCTAATTAACACAGAAATTTGGACAACTCCAGTGAAGTTCTTCT 2220
QY 1441 CTGGCTGGATCCGCGAGAAATGGAATGATGATACCGGTGAGTTTACCCGCGGCGCGCTT 1500
DB 2219 C---CTTTACTCATCGGTACCCAGCTTTTGTTCCTTTAGTGAGGGTTAATTTGCGGCTT 2163
QY 1501 GCGTAAATCATGTCATAGCTGTTTCTGTGTGTAATTTGTTATTCGCTCACAATTTCCCA 1560
DB 2162 GCGTAAATCATGTCATAGCTGTTTCTGTGTGTAATTTGTTATTCGCTCACAATTTCCCA 2103
QY 1561 CAACTACGAGCGGAGCATAAAGTGTAAAGCTGGGGTGCCTAATAGTGAAGTAACT 1620
DB 2102 CAACTACGAGCGGAGCATAAAGTGTAAAGCTGGGGTGCCTAATAGTGAAGTAACT 2043
QY 1621 CACATTAATTTGCTTTCGCTCATCTCCGCTTTTCCAGTCGGGAACCTGTCGTGCCAGCT 1680
DB 2042 CACATTAATTTGCTTTCGCTCATCTCCGCTTTTCCAGTCGGGAACCTGTCGTGCCAGCT 1983
QY 1681 GCATTAATGAAATCGGCAACCGCGGGAGAGCGGTTTGGCTATTGGGCGCTTTCGCC 1740
DB 1982 GCATTAATGAAATCGGCAACCGCGGGAGAGCGGTTTGGCTATTGGGCGCTTTCGCC 1923
QY 1741 TTTCTCGCTCACTGACTCGCTCGCTCGGTTCGTTTCGCTGCGGCGAGCGGTATCAGCTCA 1800
DB 1922 TTTCTCGCTCACTGACTCGCTCGCTCGGTTCGTTTCGCTGCGGCGAGCGGTATCAGCTCA 1863
QY 1801 CTCAAAGCGGTAAATACCGTTATTCACAGAAATCAGGGGATAACGCGAGGAAGAAATGTG 1860
DB 1862 CTCAAAGCGGTAAATACCGTTATTCACAGAAATCAGGGGATAACGCGAGGAAGAAATGTG 1803
QY 1861 AGCAAAAGCGCAGCAAAAGGCCAGGAACCGTAAAGAGCGCGGTTGCTGCGGTTTTC 1920
DB 1802 AGCAAAAGCGCAGCAAAAGGCCAGGAACCGTAAAGAGCGCGGTTGCTGCGGTTTTC 1743
QY 1921 TAGGCTCGCCCCCTTGACGAGCATCAGAAAATCGACCTCAAGTCAGAGGTGGCGAAA 1980
DB 1742 TAGGCTCGCCCCCTTGACGAGCATCAGAAAATCGACCTCAAGTCAGAGGTGGCGAAA 1683
```

QY	1981	CCCGCAGGACTATAAGATACACGGCGTTTCCCTCGTGAAGCTCCCTCGTGGCTCTCC	2040	Db	602	AGTTACATGATCCCCCATGTTGTGCAAAAAGCGGTAGTCTCCTTCGGTCTCCGATCGT	543
Db	1682	CCCGACAGACTATAAAGATACACGGCGTTTCCCTCGTGAAGCTCCCTCGTGGCTCTCC	1623	QY	3121	TGTCAGAAAGTAGTTGGCGCAGTGTATTCACTCATGTTGATGGCAGCACTGCATAATTC	3180
QY	2041	TGTTCCGACCTCGCGCTTACCGGATACTGTCCGCTTTTCTCCTTCCGGGAAGCGTGGC	2100	Db	542	TGTCAGAAAGTAGTTGGCGCAGTGTATTCACTCATGTTGATGGCAGCACTGCATAATTC	483
Db	1622	TGTTCCGACCTCGCGCTTACCGGATACTGTCCGCTTTTCTCCTTCCGGGAAGCGTGGC	1563	QY	3181	TCCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAAGTACTCAACCAAGTC	3240
QY	2101	GCCTTCTCATAGCTCACGCTGTAGTATCTCATGTTCCGTTAGTGTGCTCGCTCAAGCT	2160	Db	482	TCTTACTGTATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAAGTACTCAACCAAGTC	423
Db	1562	GCCTTCTCATAGCTCACGCTGTAGTATCTCATGTTCCGTTAGTGTGCTCGCTCAAGCT	1503	QY	3241	ATTCTGAGAATAGTGTATCGCGCAGCCGAGTTGCTTTTCCCGCGGTCAATAACGGGATAA	3300
QY	2161	GGGCTGTGTGACGAACCCCGCTTCAGCGCCGACCGCTGCGCTTATCCGGTAACTATCG	2220	Db	422	ATTCTGAGAATAGTGTATCGCGCAGCCGAGTTGCTTTTCCCGCGGTCAATAACGGGATAA	363
Db	1502	GGGCTGTGTGACGAACCCCGCTTCAGCGCCGACCGCTGCGCTTATCCGGTAACTATCG	1443	QY	3301	TACCGCGCCATAGCAGAACTTTAAAGTGTCTCATCATTTGGAACACGTTCTTCGGGGCG	3360
QY	2221	TCTTGAGTCCAAACCCGGTAAGACACGACTTATGCGCACTGGCAGCAGCACTGGTAAACAG	2280	Db	362	TACCGCGCCATAGCAGAACTTTAAAGTGTCTCATCATTTGGAACACGTTCTTCGGGGCG	303
Db	1442	TCTTGAGTCCAAACCCGGTAAGACACGACTTATGCGCACTGGCAGCAGCACTGGTAAACAG	1383	QY	3361	AAAACCTCTCAAGGATCTTTACCGCTGTTGAGATCCAGTTTCGATGTAAACCCACTCGTGCACC	3420
QY	2281	GATTAGCAGCGAGTATGTAGCGGTGCTACAGAGTTCTTTGAAGTGGTGGCCCTAACTA	2340	Db	302	AAAACCTCTCAAGGATCTTTACCGCTGTTGAGATCCAGTTTCGATGTAAACCCACTCGTGCACC	243
Db	1382	GATTAGCAGCGAGTATGTAGCGGTGCTACAGAGTTCTTTGAAGTGGTGGCCCTAACTA	1323	QY	3421	CAACTGATCTTCAGCATCTTTTACTTTTCCAGCGCTTTCTGGTGAGCAAAAACAGGAAG	3480
QY	2341	CGGCTACACTAGAAGCAGATATTTGGTATCTGCGCTCTGCTGAAGCCAGTTTACCTTCGG	2400	Db	242	CAACTGATCTTCAGCATCTTTTACTTTTCCAGCGCTTTCTGGTGAGCAAAAACAGGAAG	183
Db	1322	CGGCTACACTAGAAGCAGATATTTGGTATCTGCGCTCTGCTGAAGCCAGTTTACCTTCGG	1263	QY	3481	GCAAAATCCGCAAAAAGGGAATTAAGGCGACACGGAATGTTGAATACTCATCTCTT	3540
QY	2401	AAAAAGAGTTGATGCTTGTATCCGGCAAAACAAACCCCGCTGGTAGCGGTGTTTTT	2460	Db	182	GCAAAATCCGCAAAAAGGGAATTAAGGCGACACGGAATGTTGAATACTCATCTCTT	123
Db	1262	AAAAAGAGTTGATGCTTGTATCCGGCAAAACAAACCCCGCTGGTAGCGGTGTTTTT	1203	QY	3541	CCTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATT	3600
QY	2461	TGTTTGCAGCAGCAGATTTACCGCAGAAAAAGATCTCAAGAGATCTCTTTCATCTT	2520	Db	122	CCTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATT	63
Db	1202	TGTTTGCAGCAGCAGATTTACCGCAGAAAAAGATCTCAAGAGATCTCTTTCATCTT	1143	QY	3601	TCAATGTTATTTAGAAAAATAAAACAAATAGGGGTTCCGCGCACATTTCCCGCAAAAGTGCC	3660
QY	2521	TTCTACGGGGTCTGACGCTCAGTGGAAACGAAACTCAGCTTAAGGGATTTTGGTCATGAG	2580	Db	62	TGAATGTTATTTAGAAAAATAAAACAAATAGGGGTTCCGCGCACATTTCCCGCAAAAGTGCC	3
Db	1142	TTCTACGGGGTCTGACGCTCAGTGGAAACGAAACTCAGCTTAAGGGATTTTGGTCATGAG	1083	QY	3661	AC 3662	
QY	2581	ATTATCAAAAAGGATCTTCACTAGATCTTTTAAATTTAAATGAAGTTTAAATCAAT	2640	Db		2 AC 1	
Db	1082	ATTATCAAAAAGGATCTTCACTAGATCTTTTAAATTTAAATGAAGTTTAAATCAAT	1023	RESULT 9			
QY	2641	CTAAGATATATGATGATAAATTGTTCTGACAGTTTACCAATCTTAATCAGTGAGGCACC	2700	ADQ88423/c			
Db	1022	CTAAGATATATGATGATAAATTGTTCTGACAGTTTACCAATCTTAATCAGTGAGGCACC	963	ID	ADQ88423 standard; DNA; 3637 BP.		
QY	2701	TATCTCAGCGATCTGTCTATTTCGTTTCATCCATAGTTGCTGACTCCCGCTCGTGTAGAT	2760	XX	ADQ88423;		
Db	962	TATCTCAGCGATCTGTCTATTTCGTTTCATCCATAGTTGCTGACTCCCGCTCGTGTAGAT	903	AC	18-NOV-2004 (first entry)		
QY	2761	AACTACGATACGGGAGGGCTTACCATTGGCCCCAGTGTGCAATGATACCGCGAGACCC	2820	DT			
Db	902	AACTACGATACGGGAGGGCTTACCATTGGCCCCAGTGTGCAATGATACCGCGAGACCC	843	XX	PBSC3GFP DNA encoding Jellyfish cycle 3 GFP protein.		
QY	2821	AGCTCACCGGCTCCAGATTTATCAGCAATAAACAGCCAGCGGAAGGGCCGCGGAG	2880	DE	Molecular biology; jellyfish; green fluorescent protein; GFP; ds.		
Db	842	AGCTCACCGGCTCCAGATTTATCAGCAATAAACAGCCAGCGGAAGGGCCGCGGAG	783	KW	Aequorea victoria.		
QY	2881	AAAGTGGTCTGCAAACTTTATCCGCTCCATCCAGTCTATTATTTGTTGCGGGAAGCTAG	2940	OS	US2004142433-Al.		
Db	782	AAAGTGGTCTGCAAACTTTATCCGCTCCATCCAGTCTATTATTTGTTGCGGGAAGCTAG	723	PN	22-JUL-2004.		
QY	2941	AGTAAGTAGTTCGCGAGTTAATAGTTTTCGCAACGTTGTCATTTGTCATGACAGGCAATCGT	3000	PD	10-OCT-2003; 2003US-00684134.		
Db	722	AGTAAGTAGTTCGCGAGTTAATAGTTTTCGCAACGTTGTCATTTGTCATGACAGGCAATCGT	663	PF			
QY	3001	GGTGTACGCTCGTGGTTGGTATGGCTCATTCAGCTCCGGTTCCTCAACGATCAAGGGC	3060	XX	02-FEB-2001; 2001US-0266386P.		
Db	662	GGTGTACGCTCGTGGTTGGTATGGCTCATTCAGCTCCGGTTCCTCAACGATCAAGGGC	603	PR	14-FEB-2001; 2001US-0268785P.		
QY	3061	AGTTACATGATCCCCCATGTTGTGCAAAAAGCGGTAGTCTCCTTCGGTCTCCGATCGT	3120	PR	01-FEB-2002; 2002US-00066390.		
				PR	08-AUG-2002; 2002US-0402342P.		
				PR	21-AUG-2002; 2002US-00226372.		
				PR	25-OCT-2002; 2002US-00280913.		
				XX	08-AUG-2003; 2003US-00637758.		
				PA	(PADG/) PADGETT H S.		

Db 2042 CACATTAATTGGCTGCTCACTGCCGCTTCCAGTCGGGAACCTGCTGCGCAGCT 1983
QY 1681 GCATTAATAAATCGGCCAAGCGCGGGAGAGCGGTTTGGCTATTTGGCGGCTCTTCCG 1740
Db 1982 GCATTAATAAATCGGCCAAGCGCGGGAGAGCGGTTTGGCTATTTGGCGGCTCTTCCG 1923
QY 1741 TTCTCTCGCTCACTGACTCGCTCGGCTCGGTCGGTTCGGCTCGCGGAGCGGTATCAGCTCA 1800
Db 1922 TTCTCTCGCTCACTGACTCGCTCGGTCGGTTCGGTTCGGCTCGCGGAGCGGTATCAGCTCA 1863
QY 1801 CTCAAAGCGGTAAATACGGTTATCCACAGAAATCAGGGGATTAACGAGGAAAGAAATGTG 1860
Db 1862 CTCAAAGCGGTAAATACGGTTATCCACAGAAATCAGGGGATTAACGAGGAAAGAAATGTG 1803
QY 1861 AGCAAAAGCGCCAGCAAAAGGCCAGGAAACCGTAAAGAGCGCGTTCGTCGGCGTTTTC 1920
Db 1802 AGCAAAAGCGCCAGCAAAAGGCCAGGAAACCGTAAAGAGCGCGTTCGTCGGCGTTTTC 1743
QY 1921 TAGGCTCGCGCCCTCGACGAGCATCAAAAATCGAAGCTCAAGTCAGAGGTGCGGAAA 1980
Db 1742 TAGGCTCGCGCCCTCGACGAGCATCAAAAATCGAAGCTCAAGTCAGAGGTGCGGAAA 1683
QY 1981 CCCGACAGACTATATAAGATACAGCGGTTTCCCGCTGGAAGCTCCCTCGTCGCTCTCC 2040
Db 1682 CCCGACAGACTATATAAGATACAGCGGTTTCCCGCTGGAAGCTCCCTCGTCGCTCTCC 1623
QY 2041 TGTTCCGACCCCTGCGCTTACCGGATACCTGTCGCGCTTTCCTTCCGGAAGCGTGGC 2100
Db 1622 TGTTCCGACCCCTGCGCTTACCGGATACCTGTCGCGCTTTCCTTCCGGAAGCGTGGC 1563
QY 2101 GCTTTCTCATAGCTCACTGCTGAGTATCTCAGTTTCGGTGTAGGTTCGTTTCGCTCAAGCT 2160
Db 1562 GCTTTCTCATAGCTCACTGCTGAGTATCTCAGTTTCGGTGTAGGTTCGTTTCGCTCAAGCT 1503
QY 2161 GGGCTGTGTGACGAAACCCCGCTTACGCCGACCGCTGCGCTTATCCGGTAACTATCG 2220
Db 1502 GGGCTGTGTGACGAAACCCCGCTTACGCCGACCGCTGCGCTTATCCGGTAACTATCG 1443
QY 2221 TCTTCAGTCCACCGGTAAAGACAGACTTATCGCCACTTGGCAGCAGCCACTGTGTAAACAG 2280
Db 1442 TCTTCAGTCCACCGGTAAAGACAGACTTATCGCCACTTGGCAGCAGCCACTGTGTAAACAG 1383
QY 2281 GATTAGCAGAGCGAGTATGTAGGCGGTCTACAGAGTTCTTGAAGTGTGGCTTAACTA 2340
Db 1382 GATTAGCAGAGCGAGTATGTAGGCGGTCTACAGAGTTCTTGAAGTGTGGCTTAACTA 1323
QY 2341 CGGCTACACTAGAAAGGACAGTATTTGGTATCTGCGCTCTGCTGGAAGCGAGTTACCTTCGG 2400
Db 1322 CGGCTACACTAGAAAGGACAGTATTTGGTATCTGCGCTCTGCTGGAAGCGAGTTACCTTCGG 1263
QY 2401 AAAAGAGTTGGTAGCTCTTGATCCGGCAACAAACACCGCTGTAGGGTGGTTTTTT 2460
Db 1262 AAAAGAGTTGGTAGCTCTTGATCCGGCAACAAACACCGCTGTAGGGTGGTTTTTT 1203
QY 2461 TGTTTTCGAGCAGCAGATTACCGCAGAAAAAAGGATCTCAAGAAGATCTCTTTCATCTT 2520
Db 1202 TGTTTTCGAGCAGCAGATTACCGCAGAAAAAAGGATCTCAAGAAGATCTCTTTCATCTT 1143
QY 2521 TTCTACGGGCTGACGCTCAGTGTGAACGAAACCTCAGCTTAAAGGATTTGGTTCATGAG 2580
Db 1142 TTCTACGGGCTGACGCTCAGTGTGAACGAAACCTCAGCTTAAAGGATTTGGTTCATGAG 1083
QY 2581 ATTATCAAAAAGGATCTTCACTTACCTTAAATTAATAATCAAGTTTAAATCAAT 2640
Db 1082 ATTATCAAAAAGGATCTTCACTTACCTTAAATTAATAATCAAGTTTAAATCAAT 1023
QY 2641 CTAAAGTATATAGTAAACTTTGGTCTGACAGTTTACCAATGTTAAATGAGTGGCACC 2700
Db 1022 CTAAAGTATATAGTAAACTTTGGTCTGACAGTTTACCAATGTTAAATGAGTGGCACC 963
QY 2701 TATCTCAGCGATCTGCTATTTGTTGTTTCATTCATAGTTCGCTGACCTCCCGTCTGTAGAT 2760
|||||

Db 962 TATCTCAGCGATCTGCTCTATTTTGGTTTCATCCATAGTTCGCTGACCTCCCGTCTGTAGAT 903
QY 2761 AACTACGATACGGGAGGCTTACCATCTGGCCCCAGTGTGCAATGATATACCCGAGACCC 2820
Db 902 AACTACGATACGGGAGGCTTACCATCTGGCCCCAGTGTGCAANTGATACCCGAGACCC 843
QY 2821 AGCTCACCGGCTCAGATTTTATCAGCAATAAACAGCAGCCGGAAGGCCGAGCGCAG 2880
Db 842 AGCTCACCGGCTCAGATTTTATCAGCAATAAACAGCAGCCGGAAGGCCGAGCGCAG 783
QY 2881 AGTGGTCTGCAACTTTATCCGCCCTCCATCCAGTCTATTAATTTGTTGCCGGGAAGCTAG 2940
Db 782 AGTGGTCTGCAACTTTATCCGCCCTCCATCCAGTCTATTAATTTGTTGCCGGGAAGCTAG 723
QY 2941 AGTAAGTAGTTTCGCCAGTTAATAGTTTGCAGAACTGTTGTCATTTGTCATTTGCTACAGGCATCGT 3000
Db 722 AGTAAGTAGTTTCGCCAGTTAATAGTTTGGGCAACGTTGTTGCCATTTGCTACAGGCATCGT 663
QY 3001 GGTGTACGCTCGTTCGTTTGGTATGCTTTCATTTACAGTCCGGTTCCTCAACGATCAAGGCG 3060
Db 662 GGTGTACGCTCGTTCGTTTGGTATGGCTTCATTTACAGCTCCGGTTCCTCAACGATCAAGGCG 603
QY 3061 AGTTACATGATCCCCCATGTTGTGCAAAAAGCGGTTAGCTCCTTCGGTCTCCGATCGT 3120
Db 602 AGTTACATGATCCCCCATGTTGTGCAAAAAGCGGTTAGCTCCTTCGGTCTCCGATCGT 543
QY 3121 TGTCAAGAGTAAGTTGGCGCAGTGTATCACTCATGTTATGGCAGCAGCTCATATAATTC 3180
Db 542 TGTCAAGAGTAAGTTGGCGCAGTGTATCACTCATGTTATGGCAGCAGCTCATATAATTC 483
QY 3181 TCTTACTGTCATGTCGTAAGATGCTTTTCTGTGACTGGTGTAGTACTCAACCAAGTC 3240
Db 482 TCTTACTGTCATGTCGTAAGATGCTTTTCTGTGACTGGTGTAGTACTCAACCAAGTC 423
QY 3241 ATTCTGAGAAATAGTGTATGGCGGACCGAGTGTCTTTGCCGGCGGTCAATACGGGATTA 3300
Db 422 ATTCTGAGAAATAGTGTATGGCGGACCGAGTGTCTTTGCCGGCGGTCAATACGGGATTA 363
QY 3301 TACCGCGCACATAGCAGAACTTTAAAAGTGTCACTCATTTGGAAAACTTTCTTCGGGGCG 3360
Db 362 TACCGCGCACATAGCAGAACTTTAAAAGTGTCACTCATTTGGAAAACTTTCTTCGGGGCG 303
QY 3361 AAACTCTCAAGGATCTTTACCGCTGTGTGAGATCCAGTTCGATGTAAACCCACTTCGTGCACC 3420
Db 302 AAACTCTCAAGGATCTTTACCGCTGTGTGAGATCCAGTTCGATGTAAACCCACTTCGTGCACC 243
QY 3421 CAACTGATCTTCAGCATCTTTTACTTTTCCAGCGGTTTCTGGGTGAGCAAAAACAGGAG 3480
Db 242 CAACTGATCTTCAGCATCTTTTACTTTTCCAGCGGTTTCTGGGTGAGCAAAAACAGGAG 183
QY 3481 GCAAAATCCGCAAAAAGGGAATAAGGGCGACACGGAATAAGTTGAATACTCATACTCTT 3540
Db 182 GCAAAATCCGCAAAAAGGGAATAAGGGCGACACGGAATAAGTTGAATACTCATACTCTT 123
QY 3541 CCTTTTCAATATTAATGAAGCATTTATCAGGGTTATTTGTTCTCATGAGCGGATACATATT 3600
Db 122 CCTTTTCAATATTAATGAAGCATTTATCAGGGTTATTTGTTCTCATGAGCGGATACATATT 63
QY 3601 TGAATGATTTAGAAAAATAAATAAGGGGTTCCGCGCACATTTCCCGCAAAAGTGCC 3660
Db 62 TGAATGATTTAGAAAAATAAATAAGGGGTTCCGCGCACATTTCCCGCAAAAGTGCC 3
QY 3661 AC 3662
|||
Db 2 AC 1

RESULT 10

ADSL1717/c

ID ADSL1717 standard; DNA; 3637 BP.

XX ADSL1717;

AC ADSL1717;

XX

```
DT 02-DEC-2004 (first entry)
XX pBSC3BFP plasmid DNA.
DE
XX Polymerase; cyclic; circular; ds.
XX
XX Unidentified.
XX
XX US2004180352-A1.
XX
XX 16-SEP-2004.
XX
XX 08-AUG-2003; 2003US-00637758.
XX
XX 08-AUG-2002; 2002US-0402342P.
XX
XX (LARG-) LARGE SCALE BIOLOGY CORP.
XX
XX Padgett HS, Lindbo JA, Fitzmaurice WP, Vaewhongs AA;
XX WPI; 2004-667658/65.
XX
XX In vitro method of making sequence variants from heteroduplex
XX polynucleotide, involves combining polynucleotide with agent having
XX polymerase and strand cleavage activities for sufficient time for
XX percentage of complementarity to increase.
XX
XX Example 14; SEQ ID NO 17; 81pp; English.
XX
XX The invention relates to an in vitro method of making sequence variants
XX from heteroduplex polynucleotide which involves combining polynucleotide
XX with an agent having polymerase and strand cleavage activities for
XX sufficient time for the percentage of complementarity to be increased
XX within the heteroduplex. The method is useful for performing in vitro
XX method of making sequence variants from one or more heteroduplex
XX polynucleotide. It is useful in increasing diversity in a population of
XX sequences and obtaining a polynucleotide encoding a desired functional
XX property. The present sequence is a pBSC3BFP plasmid DNA. This sequence
XX is used to illustrate the method of the invention.
XX
XX Sequence 3637 BP; 942 A; 892 C; 870 G; 933 T; 0 U; 0 Other;
XX
XX Query Match 72.8%; Score 2664.8; DB 13; Length 3637;
XX Best Local Similarity 83.8%; Pred. No. 0;
XX Matches 3070; Conservative 0; Mismatches 567; Indels 25; Gaps 4;
XX
XX 1 CTAATTTGTAAGCGTTAATTTTGTAAATTCGGTAAATTTTGTAAATTCAGCTC 60
XX 3637 CTAATTTGTAAGCGTTAATTTTGTAAATTCGGTAAATTTTGTAAATTCAGCTC 3578
XX
XX 61 ATTTTAAACCAATAGCGGAAATCGGCAAAATCCCTTATAAATCAAAAGAAATAGACCGA 120
XX 3577 ATTTTAAACCAATAGCGGAAATCGGCAAAATCCCTTATAAATCAAAAGAAATAGACCGA 3518
XX
XX 121 GATAGGGTTGAGTGTGTTCCAGTTTGGAAACAAGAGTCCAATTAAGAAACGTGGACTC 180
XX 3517 GATAGGGTTGAGTGTGTTCCAGTTTGGAAACAAGAGTCCAATTAAGAAACGTGGACTC 3458
XX
XX 181 CACGTCGAAGGCGGAAACCCCTCTATCAGGCGATGGCCCACTACGTGAACCATCAC 240
XX 3457 CACGTCGAAGGCGGAAACCCCTCTATCAGGCGATGGCCCACTACGTGAACCATCAC 3398
XX
XX 241 CTAATCAAGTTTTTTGGGTCGAGGTGCCGTAAGCACTAAATCGGAACCCCTAAAGGGAG 300
XX 3397 CTAATCAAGTTTTTTGGGTCGAGGTGCCGTAAGCACTAAATCGGAACCCCTAAAGGGAG 3338
XX
XX 301 CCCCCGATTTAGAGTTGACGGGGAACCGCGGAACAGTGGCGGAGAAAGGAAGGAAGAA 360
XX 3337 CCCCCGATTTAGAGTTGACGGGGAACCGCGGAACAGTGGCGGAGAAAGGAAGGAAGAA 3278
XX
XX 361 AGCGAAGGACGGGCGCTAGGCGGCTGGCAAGTGTAGCGGTACCGTGGCGGTAAACCAC 420
XX 3277 AGCGAAGGACGGGCGCTAGGCGGCTGGCAAGTGTAGCGGTACCGTGGCGGTAAACCAC 3218
```

```
QY 421 CACACCCGCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTTCAGGCTGCG 480
DB |||||
DB CACACCCGCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTTCAGGCTGCG 3158
QY 481 CAACTGTTGGGAAGGGCGGATCGGTGGGGCTCTTCGCTATTACGCCAGCTGGCGGAAGG 540
DB |||||
DB CAACTGTTGGGAAGGGCGGATCGGTGGGGCTCTTCGCTATTACGCCAGCTGGCGGAAGG 3098
QY 541 GGGATGTCGTCGAAGGCGGATTAAGTTGGGTAAGCCAGGCGTTCCTCCAGTCAGACGCTTG 600
DB |||||
DB GGGATGTCGTCGAAGGCGGATTAAGTTGGGTAAGCCAGGCGTTCCTCCAGTCAGACGCTTG 3038
QY 601 TAAACGACGCGCCAGTGAGCGCGCTCGTTCAATTCAGTTTTCAGTAAACCCCTGGAGGACGG 660
DB |||||
DB TAAACGACGCGCCAGTGAGCGCGCTCGTTCAATTCAGTTTTCAGTAAACCCCTGGAGTCCA 2978
QY 661 GCAGACTCGCGGTGCAAAATGTTTACAGCGGTGATGGAGCAGATGAAGATGCTCGACAC 720
DB |||||
DB CGCGGTGGCGG-----CGCTCTAGAACTAGTGGATCCCC--CGG 2939
QY 721 GCTGCAGAAACACGCGACTAGATTAAACCCCTAGAAAGATAATCATATTGTGACGTACGTAA 780
DB |||||
DB GCTGCAGGAAATCTTATTGTTAGTTTCATCCATGCGATGTTAAATCCCGACGACGATTA 2879
QY 781 AGATAATCATGCGTAAATTTGACGATGGGATCTGTAATACGATCACTATAGGGCGAAT 840
DB |||||
DB CAAACTCAAGAGAGGACCATGTGGTCAAGCTTTTCTGTTGGGATCTTTCGAAAGGCGAGATT 2819
QY 841 TGGGTACCGGCGGCCCCCTCGAGGTGCGCGGTATCGATAGCTTGTATCGAATTCCTGC 900
DB |||||
DB GTGTCGACAGGTAATGTTGTTGTTAAAGGACAGGGCGCATCGCAATTTGGAGTATTTT 2759
QY 901 AGCCCGGGGATCCACTAGTTCTAGAGCGCGCGCCACCGCGGTGGAGCTCCAGCTTTGT 960
DB |||||
DB GTTGATAATGGTCTGCTAGTTGAAACGATCCATCTTCAATGTTGTGGGAAATTTGAAGT 2699
QY 961 TCCCTTTAGTAGGGTTAATTAGATCCCATGGTCAATTTTACGAGACTATCTTTCTAG 1020
DB |||||
DB TAGCTTTGATTCATCTTTTGTCTGCGCGGTATGATATATATCATTTGTGTGAG-TTATAG 2640
QY 1021 GGTAACTAGCTGCATCAGGATCATATCGTGGGCTTTTTCGGCTCAGTCACTCGCC 1080
DB |||||
DB TTGTACTCAGTTTGTGTCGAGAAATGTTTCCATCTTCTTTAAATAATCAATCTTTTAAAC 2580
QY 1081 CAACTGCGCGCTATCTGGGCATCGGGGAGGAAGACCGGCTTTTCCCGGAGGTG 1140
DB |||||
DB TCGATACGATTAAAGGGTATCACCTTCNAATTCAGCTCAGCAGCGCTCTTGTAGTTC 2520
QY 1141 AAGCGCATGGAAAGATTGCGGAGGATGACTGCTGCTGCTGATGACGTTGAGCGGAAAC 1200
DB |||||
DB CGTCACTTTGAAAGATATAGTGGCTTCTGTACATAACCTTCGGGCATGGCACTCTTG 2460
QY 1201 GCAGTTTACCATGATGATTCGGGAAGGTGGCCATGACGCGCTTTACGGTGAAGTGT 1260
DB |||||
DB AAAAACTCATGCGTTTTCATATGATCCGGATAACGGGAAAGCAATGAAACCATTAAGAG 2400
QY 1261 TCGTTTCAGGCGCACTCGGATACCAAGTTCTGCGGCGCTTTTCCGGAACACAGTTCCGGATGG 1320
DB |||||
DB AAGTAGTGACAGTGTGGCCATGGAAACAGGTAGTTTTCAGTAGTGCAAAATAATTTA 2340
QY 1321 TCAGCCGGAAGCGCATCAGCAACCCGAAACAATACCGCGGACACCGGAACTGCGGTGCG 1380
DB |||||
DB AGGGTAAGCTTTCCGTATGTAGCATCACCTTCACTCTCCACTGACAGAAAAATTTGTGC 2280
QY 1381 GTGTGCAGATTATGACAGGGGTGGCGCTGGGATATTACGTGAGGAGGACGGGTATC 1440
DB |||||
DB CCAATTAACATCACTAATTTAAATTTGGGACAACTCCAGTGAAAGTTCTTCT 2220
QY 1441 CTGGCTGGGATGCGCGAAGAAATGGACATGGATACCCGTGAGTTACCGGCGGCGCGCTT 1500
DB |||||
DB CTTTACTCATCGGTACCCAGCTTTTGTTCCTTTTAGTAGGGGTAAATTCGGCGCTT 2163
```

QY 1501 GGCGTAATCATGGTCATAGCTGTTCTCTGTGTGAATTTGTTATCCGCTCACAAATCCACA 1560
DB 2162 GGCGTAATCATGGTCATAGCTGTTCTCTGTGTGAATTTGTTATCCGCTCACAAATCCACA 2103
QY 1561 CAACTACGAGCCGGAAGCATAAAGTGTAAAGCTTGGGTGCTTAATGATGAGCTAACT 1620
DB 2102 CAACTACGAGCCGGAAGCATAAAGTGTAAAGCTTGGGTGCTTAATGATGAGCTAACT 2043
QY 1621 CACATTAATTTGGTGTGGCTCACTGCGCTTTCCAGTCGGGAACCTGTGTCGCCAGCT 1680
DB 2042 CACATTAATTTGGTGTGGCTCACTGCGCTTTCCAGTCGGGAACCTGTGTCGCCAGCT 1983
QY 1681 GCATTAATGAATCGGCCAAACGCGCGGGAGAGCGGTTTTCGATTTGGCGCTCTTCCG 1740
DB 1982 GCATTAATGAATCGGCCAAACGCGCGGGAGAGCGGTTTTCGATTTGGCGCTCTTCCG 1923
QY 1741 TTCTCTGCTCACTGACTCGCTCGGCTCGGTGCTTTCGGCTTGGCGAGCGGTATCAGCTCA 1800
DB 1922 TTCTCTGCTCACTGACTCGCTCGGCTCGGTGCTTTCGGCTTGGCGAGCGGTATCAGCTCA 1863
QY 1801 CTCAAAGGCGGTAAATACGGTTTATCCAGAAATCAGGGGATAACGCAAGGAAGAAATGTG 1860
DB 1862 CTCAAAGGCGGTAAATACGGTTTATCCAGAAATCAGGGGATAACGCAAGGAAGAAATGTG 1803
QY 1861 AGCAAAAGGCCACGCAAAAGGCCAGGAACCGTTAAAGAGCGCGGTTGCTGGCGTTTTC 1920
DB 1802 AGCAAAAGGCCACGCAAAAGGCCAGGAACCGTTAAAGAGCGCGGTTGCTGGCGTTTTC 1743
QY 1921 TAGGCTCGCGCCCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGTGGGGAAA 1980
DB 1742 TAGGCTCGCGCCCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGTGGGGAAA 1683
QY 1981 CCGCAGAGACTATAAAGATACAGCGGTTTCCCTGGAAGCTCCCTCGTCGCTCTCC 2040
DB 1682 CCGCAGAGACTATAAAGATACAGCGGTTTCCCTGGAAGCTCCCTCGTCGCTCTCC 1623
QY 2041 TGTTCCGACCTTCCGCTTACCGGATACCTGTCCGCTTTTCCCTTCGGGAAGCGTGGC 2100
DB 1622 TGTTCCGACCTTCCGCTTACCGGATACCTGTCCGCTTTTCCCTTCGGGAAGCGTGGC 1563
QY 2101 GCTTTTCTCATAGCTCACGCTGTAGTATCTCAGTTTCGGTGTAGTTCGTTCCAGCT 2160
DB 1562 GCTTTTCTCATAGCTCACGCTGTAGTATCTCAGTTTCGGTGTAGTTCGTTCCAGCT 1503
QY 2161 GGGCTGTGTGCAGGAACCCCGCTTACGCCGACCGCTGCGCTTATCCGGTAACTATCG 2220
DB 1502 GGGCTGTGTGCAGGAACCCCGCTTACGCCGACCGCTGCGCTTATCCGGTAACTATCG 1443
QY 2221 TCTTGAGTCCAAACCCGGTAAGACAAGACTTATCGCACTGGCAGCAGCACTGGTAAACAG 2280
DB 1442 TCTTGAGTCCAAACCCGGTAAGACAAGACTTATCGCACTGGCAGCAGCACTGGTAAACAG 1383
QY 2281 GATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTTAACTA 2340
DB 1382 GATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTTAACTA 1323
QY 2341 CGGCTACACTAGAAGCAGAGTATTTGGTATCTGCGCTCTGCTGGAAGCAAGTTACCTTCGG 2400
DB 1322 CGGCTACACTAGAAGCAGAGTATTTGGTATCTGCGCTCTGCTGGAAGCAAGTTACCTTCGG 1263
QY 2401 AAAAAAGTTGGTAGCTCTTTGATCCGGCAAAACAAACCCGCTGGTAGCGGTGGTTTTT 2460
DB 1262 AAAAAAGTTGGTAGCTCTTTGATCCGGCAAAACAAACCCGCTGGTAGCGGTGGTTTTT 1203
QY 2461 TGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAGATCTTTGATCTT 2520
DB 1202 TGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAGATCTTTGATCTT 1143
QY 2521 TTCTACGGGCTGACGCTCAGTGGAAAGAAACTCAGTTAAGGATTTTGGTCATGAG 2580
DB 1142 TTCTACGGGCTGACGCTCAGTGGAAAGAAACTCAGTTAAGGATTTTGGTCATGAG 1083
QY 2581 ATTATCAAAAAGGATCTTCACTACCTAGATCCTTTTAAATTAATAAGGATTTTAAATCAAT 2640

DB 1082 ATTATCAAAAAGGATCTTCACTACCTAGATCCTTTAAATTAATAAGTTTTAAATCAAT 1023
QY 2641 CTAAAGTATATATGATGATAAACTTGGTCTGACAGTTTACCAATCTTAATCAGTGAGGCACC 2700
DB 1022 CTAAAGTATATATGATGATAAACTTGGTCTGACAGTTTACCAATCTTAATCAGTGAGGCACC 963
QY 2701 TATCTCAGCGATCTCTGCTATTTTGGTTCATCCATAGTTGCTGACTCCCCCGTGGTAGAT 2760
DB 962 TATCTCAGCGATCTCTGCTATTTTGGTTCATCCATAGTTGCTGACTCCCCCGTGGTAGAT 903
QY 2761 AACTACGATACCGGAGGGCTTACCATCTGCGCCCCAGTGTGCAATGATACCGGAGACCC 2820
DB 902 AACTACGATACCGGAGGGCTTACCATCTGCGCCCCAGTGTGCAATGATACCGGAGACCC 843
QY 2821 AGGCTCACCGGCTCCAGATTTATCAGCAATAAACAGCCAGCCGGAAGGGCGAGGCGAG 2880
DB 842 AGGCTCACCGGCTCCAGATTTATCAGCAATAAACAGCCAGCCGGAAGGGCGAGGCGAG 783
QY 2881 AAGTGGTCTGCAACTTTATCCGCTCCATCCAGTCTATTAAATTTGTCGGGAAGCTAG 2940
DB 782 AAGTGGTCTGCAACTTTATCCGCTCCATCCAGTCTATTAAATTTGTCGGGAAGCTAG 723
QY 2941 AGTAAAGTATTGCGCAGTTAATAGTTTGGCAACGTTGTTGCCATTTGCTACAGGCATCT 3000
DB 722 AGTAAAGTATTGCGCAGTTAATAGTTTGGCAACGTTGTTGCCATTTGCTACAGGCATCT 663
QY 3001 GGTGTACGCTCGTGGTGTGATGCTTCAATTCAGCTCCGTTCCCAACGATCAAGGG 3060
DB 662 GGTGTACGCTCGTGGTGTGATGCTTCAATTCAGCTCCGTTCCCAACGATCAAGGG 603
QY 3061 AGTTACATGATCCCCCATGTTGTGCAAAAAGCGGTAGCTCCTTCGGTCTCCGATCT 3120
DB 602 AGTTACATGATCCCCCATGTTGTGCAAAAAGCGGTAGCTCCTTCGGTCTCCGATCT 543
QY 3121 TGTCAAGAGTAAGTTGGCGCAGTGTATCACTCATGTTTATGGCAGCAGCTCATAAATTC 3180
DB 542 TGTCAAGAGTAAGTTGGCGCAGTGTATCACTCATGTTTATGGCAGCAGCTCATAAATTC 483
QY 3181 TCTTACTGTCACTGCTCCGTAAGATGCTTTTCTGCTGACTGGTGTAGTCTCAACCAAGTC 3240
DB 482 TCTTACTGTCACTGCTCCGTAAGATGCTTTTCTGCTGACTGGTGTAGTCTCAACCAAGTC 423
QY 3241 ATTCTGAGAATAGTGTATGTCGGCGACCGAGTTGCTCTTTCGCCGCGTCAATACGGGATAA 3300
DB 422 ATTCTGAGAATAGTGTATGTCGGCGACCGAGTTGCTCTTTCGCCGCGTCAATACGGGATAA 363
QY 3301 TACCGCGCCACATAGCAGAACTTTAAAGTGTCTCATCATTTGGAAAAAGTCTTTCGGGGCG 3360
DB 362 TACCGCGCCACATAGCAGAACTTTAAAGTGTCTCATCATTTGGAAAAAGTCTTTCGGGGCG 303
QY 3361 AAACTCTCAAGGATCTTACCGCTGTGAGATCCAGTTTCGATTAACCCACCTCGTGCACC 3420
DB 302 AAACTCTCAAGGATCTTACCGCTGTGAGATCCAGTTTCGATTAACCCACCTCGTGCACC 243
QY 3421 CAACTGATCTTCAGCATCTTTTACTTTTCCAGCGCTTTCTGGGTGAGCAAAAACAGGAAG 3480
DB 242 CAACTGATCTTCAGCATCTTTTACTTTTCCAGCGTCTTTCGGGTGAGCAAAAACAGGAAG 183
QY 3481 GAAAAATCCGCAAAAAAGGGAATTAAGGGCAGACGGAATAATGTTGAATACTCATACTCTT 3540
DB 182 GAAAAATCCGCAAAAAAGGGAATTAAGGGCAGACGGAATAATGTTGAATACTCATACTCTT 123
QY 3541 CCTTTTCAATATTTATCAAGCATTTATCAGGTTATTGCTCTCATGAGGGATACATATT 3600
DB 122 CCTTTTCAATATTTATCAAGCATTTATCAGGTTATTGCTCTCATGAGGGATACATATT 63
QY 3601 TGAATGATTTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCGAAAAAGTGCC 3660
DB 62 TGAATGATTTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCGAAAAAGTGCC 3
QY 3661 AC 3662

Db 2394 AGTGCAAGTGTGGCCATGGAACAGGTAGTATTTCCAGTAGTGCAAAATAAATTTAAGGGT 2335
QY 1319 GGTGACGCCGAAGCGCATCAGCAACCCGAAACAATACCGCGGACAGCCGGAACCTGCGGTGC 1378
Db 2334 AAGTTTTC-----CGTATGTTGCATCACCTTCCACCTCTCCACTGACAGAAAATTTGTGC 2280
QY 1379 CGGTGTGCAGATTAATGACAGCGGTGCGCGCTGGGATATTACGTGACGAGGAGCGGGTA 1438
Db 2279 CCAATTAACATCACCATTAAATCAACAGAAATTTGGGACAACTCCAGTGAAAGTCTTCT 2220
QY 1439 TCTGCGTGGATGCGCGCAGAAATGGACATGGAATACCCCTGAGTTACCGCGCGGCGCGC 1498
Db 2219 CTTTACTCATCGGTACCCAGCTTTTGTTCCTTTAGTGAGGGTTA-----ATTGCGGC 2165
QY 1499 TTGGCGTAATCATGCTCATAGCTGTTTCTGTGTGAAATGTTATCCGCTCAATTTCA 1558
Db 2164 TTGGCGTAATCATGCTCATAGCTGTTTCTGTGTGAAATGTTATCCGCTCAATTTCA 2105
QY 1559 CACAACATACGAGCCGGAAGCAATAAGTGTAAAGCTTGGGGTGCCTAATGAGTGAGCTAA 1618
Db 2104 CACAACATACGAGCCGGAAGCAATAAGTGTAAAGCTTGGGGTGCCTAATGAGTGAGCTAA 2045
QY 1619 CTCACATTAATTTGGTGGCTCACTGCCCCCTTTCCAGTCCGGAAACCTGTGCTGCCAG 1678
Db 2044 CTCACATTAATTTGGTGGCTCACTGCCCCCTTTCCAGTCCGGAAACCTGTGCTGCCAG 1985
QY 1679 CTGCATTAATGAATCGGCCAAACGCGGGGAGAGCGGTTTGGGTATTTGGGCGCTTCC 1738
Db 1984 CTGCATTAATGAATCGGCCAAACGCGGGGAGAGCGGTTTGGGTATTTGGGCGCTTCC 1925
QY 1739 GTTCTCCTGCTCACTGACTGCTGCTGCGTTCGGCTGCGCGAGCGGTATCAGCT 1798
Db 1924 GTTCTCCTGCTCACTGACTGCTGCTGCGTTCGGCTGCGCGAGCGGTATCAGCT 1865
QY 1799 CACTCAAAAGCGGTAAATACGGTTATCCACAGAAATCAGGGGATAACGCAAGAAAGAACATG 1858
Db 1864 CACTCAAAAGCGGTAAATACGGTTATCCACAGAAATCAGGGGATAACGCAAGAAAGAACATG 1805
QY 1859 TGAGCAAAAGGCGCAGAAAGCCAGGACCGTAAGAGCCGCTTGCCTGCGCTTTTC 1918
Db 1804 TGAGCAAAAGGCGCAGAAAGGCCAGGAAACCGTAAGAGCCGCTTGCCTGCGCTTTTC 1745
QY 1919 CATTAGCTCCGCCCCCTTGACGAGCATCAAAAAATCGACGCTCAAGTCAAGTGGGCGCA 1978
Db 1744 CATTAGCTCCGCCCCCTTGACGAGCATCAAAAAATCGACGCTCAAGTCAAGTGGGCGCA 1685
QY 1979 AACCCGACAGGACTATAAGATACAGGCGTTTCCCTTGGAAAGCTCCCTCGTGGCTCT 2038
Db 1684 AACCCGACAGGACTATAAGATACAGGCGTTTCCCTTGGAAAGCTCCCTCGTGGCTCT 1625
QY 2039 CCGTTCGACCCCTGCGGCTTACCGGATACCTGTGCGGCTTCTCCCTTCGGAAGCGTG 2098
Db 1624 CCGTTCGACCCCTGCGGCTTACCGGATACCTGTGCGGCTTCTCCCTTCGGAAGCGTG 1565
QY 2099 GCGCTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTTCGCTCCAAAG 2158
Db 1564 GCGCTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTTCGCTCCAAAG 1505
QY 2159 CTGGGCTGTGTGCAAGAACCCCGCTTACGCCGACCGCTGCGCTTATTCGGGTAACTAT 2218
Db 1504 CTGGGCTGTGTGCAAGAACCCCGCTTACGCCGACCGCTGCGCTTATTCGGGTAACTAT 1445
QY 2219 CGTCTGTAGTCCAAACCGGTAGACACGACTTATCGCCACTGGCAGCGCCACTGGTAAC 2278
Db 1444 CGTCTGTAGTCCAAACCGGTAGACACGACTTATCGCCACTGGCAGCGCCACTGGTAAC 1385
QY 2279 AGGATTTAGCAGAGCAGGTATGTAGCGGTGCTACAGAGTCTCTGAAGTGGTGGCTTAAC 2338
Db 1384 AGGATTTAGCAGAGCAGGTATGTAGCGGTGCTACAGAGTCTCTGAAGTGGTGGCTTAAC 1325
QY 2339 TACGCGCTACCTAGAGGACAGTATTTGGTATCTCGCTCTGCTGAAGCCAGTTACCTTC 2398

Db 1324 TACGCGCTACCTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTC 1265
QY 2399 GGAAAAAGAGTTGGTAGCTCTTGATCCGCGCAAAACAAACACCGCTGGTAGCGGTGTTTT 2458
Db 1264 GGAAAAAGAGTTGGTAGCTCTTGATCCGCGCAAAACAAACACCGCTGGTAGCGGTGTTTT 1205
QY 2459 TTTGTTTCAAGCAGCAGATACGCGCAGAAAAAAGGATCTCAAGAAAGATCCTTTGATC 2518
Db 1204 TTTGTTTCAAGCAGCAGATACGCGCAGAAAAAAGGATCTCAAGAAAGATCCTTTGATC 1145
QY 2519 TTTTCTACGGGCTCTGACGCTCAGTGGAAACGAAACTCACTGTTAAGGATTTTGGTCTAG 2578
Db 1144 TTTTCTACGGGCTCTGACGCTCAGTGGAAACGAAACTCACTGTTAAGGATTTTGGTCTAG 1085
QY 2579 AGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTTAAATTTAAATTTAAATCA 2638
Db 1084 AGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTTAAATTTAAATTTAAATCA 1025
QY 2639 ATCTAAAGTATATATGAGTAAACCTTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGCA 2698
Db 1024 ATCTAAAGTATATATGAGTAAACCTTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGCA 965
QY 2699 CCTATCTCAGCGATCTGTCTATTTGTTTCATCCATAGTTGCTGACTCCCGTCTGTAG 2758
Db 964 CCTATCTCAGCGATCTGTCTATTTGTTTCATCCATAGTTGCTGACTCCCGTCTGTAG 905
QY 2759 ATAACTACGATACGGGAGGCTTACCATCTGCCCCAGTGTGCAATGATGATACCGGAGAC 2818
Db 904 ATAACTACGATACGGGAGGCTTACCATCTGCCCCAGTGTGCAATGATGATACCGGAGAC 845
QY 2819 CACGCTCACCGGCTCCAGATTTTATCAGCAATAAACACGAGCCAGGAGGCGCGAGCGC 2878
Db 844 CACGCTCACCGGCTCCAGATTTTATCAGCAATAAACACGAGCCAGGAGGCGCGAGCGC 785
QY 2879 AGAAGTGTCTCGAACTTTTATCCGCTCCATCCAGTCTTAAATTTGTTGCGGGAAGCT 2938
Db 784 AGAAGTGTCTCGAACTTTTATCCGCTCCATCCAGTCTTAAATTTGTTGCGGGAAGCT 725
QY 2939 AGAGTAAGTAGTTCGCGAGTTAATAGTTTGGCAACGTTTGGCCATTTGCTACAGGCATC 2998
Db 724 AGAGTAAGTAGTTCGCGAGTTAATAGTTTGGCAACGTTTGGCCATTTGCTACAGGCATC 665
QY 2999 GTGGTGTACGCTCGCTGTTTGGTATGGCTTTCAATTCAGTCCGGTTCCTCAACGATCAAG 3058
Db 664 GTGGTGTACGCTCGCTGTTTGGTATGGCTTTCAATTCAGTCCGGTTCCTCAACGATCAAG 605
QY 3059 CGAGTTAATGATCCCCCATGTTGTGCAAAAAGCGGTTAGTCTCTCGGTCTCCGATC 3118
Db 604 CGAGTTAATGATCCCCCATGTTGTGCAAAAAGCGGTTAGTCTCTCGGTCTCCGATC 545
QY 3119 GTTGTCAAGTAAGTTGCGCGCAGGTTTATCAGTCTATGTTATGCGCAGCAGTGCATAT 3178
Db 544 GTTGTCAAGTAAGTTGCGCGCAGGTTTATCAGTCTATGTTATGCGCAGCAGTGCATAT 485
QY 3179 TCTCTTACTGTCAATCGCAATCGGTAAGATGCTTTTCTGTGACTGGTGTAGTCTCAACCAAG 3238
Db 484 TCTCTTACTGTCAATCGGTAAGATGCTTTTCTGTGACTGGTGTAGTCTCAACCAAG 425
QY 3239 TCATTTCTGAGAAATAGTGTATGCGGGGACGAGTGTCTTTCGCCGCGCTCAATACGGAT 3298
Db 424 TCATTTCTGAGAAATAGTGTATGCGGGGACGAGTGTCTTTCGCCGCGCTCAATACGGAT 365
QY 3299 AATACCGGCCACATAGCAGAACTTTTAAAGTGTCTATTTGAAAAACGTTCTTCGGGG 3358
Db 364 AATACCGGCCACATAGCAGAACTTTTAAAGTGTCTATTTGAAAAACGTTCTTCGGGG 305
QY 3359 CGAAAACTCTCAAGGATCTTACCGCTGTGTAGATCCAGTTCGATGTAAACCACTCGTGCA 3418
Db 304 CGAAAACTCTCAAGGATCTTACCGCTGTGTAGATCCAGTTCGATGTAAACCACTCGTGCA 245
QY 3419 CCCAACTGATCTTCAGCATCTTTTACTTTTCAACGAGCTTTTCTGGGTGAGCAAAAAACGGA 3478
Db 244 CCCAACTGATCTTCAGCATCTTTTACTTTTCAACGAGCTTTTCTGGGTGAGCAAAAAACGGA 185

QY 3479 AGCAAAATGCCCAAAAAAGGAATAAGGCGACACGGAATGTTGAATACCTACTACTC 3538
DB 184 AGCAAAATGCCCAAAAAAGGAATAAGGCGACACGGAATGTTGAATACCTACTACTC 125
QY 3539 TTCTTTTCAATATTATTGAAGCATTTATCAGGTTATTGCTCATGAGCGGATACATA 3598
DB 124 TTCTTTTCAATATTATTGAAGCATTTATCAGGTTATTGCTCATGAGCGGATACATA 65
QY 3599 TTGGAATGTTATAGAAAAATAAAATAGGGGTTCCGCGCACATTTCCCGGAAAGTG 3658
DB 64 TTGGAATGTTATAGAAAAATAAAATAGGGGTTCCGCGCACATTTCCCGGAAAGTG 5
QY 3659 CCAC 3662
DB 4 CCAC 1

RESULT 12
ADM68462/c
ID ADM68462 standard; DNA; 3637 BP.
XX
AC ADM68462;
XX
DT 03-JUN-2004 (first entry)
XX
DE Wild type green fluorescent protein, GFP, gene construct.
XX
KW ds; mismatch endonuclease; endonuclease; gene shuffling technology;
KW single nucleotide polymorphism; cancer susceptibility;
KW sequence variation redistribution; GFP; green fluorescent protein; gene.
XX
OS Aequorea victoria.
OS Synthetic.
XX
PN US2003157682-A1.
XX
PD 21-AUG-2003.
XX
PF 31-JAN-2003; 2003US-00356708.
PR 01-FEB-2002; 2002US-0353722P.
PR 14-MAR-2002; 2002US-00098155.
PR 01-AUG-2002; 2002US-00211079.
XX
PA (PADG/) PADGETT H S.
PA (VAEW/) VAETHONGS A A.
PA (VOJD/) VOJDANI F S.
PA (SMIT/) SMITH M L.
PA (LIND/) LINDBO J A.
PA (FITZ/) FITZMAURICE W P.
XX
PI Padgett HS, Vaethongs AA, Vojdani FS, Smith ML, Lindbo JA;
PI Fitzmaurice WP;
XX
DR WPI; 2003-766176/72.
XX
PT Making a mismatch endonuclease, useful in gene shuffling and in detection
PT of single nucleotide polymorphisms, comprises transfecting a host with a
PT recombinant viral vector including a polynucleotide encoding a mismatch
PT endonuclease.
XX
PS Example 5; SEQ ID NO 31; 79pp; English.
XX
CC The invention relates to a method of making a mismatch endonuclease
CC enzyme comprising transfecting a host plant, animal, yeast, fungus or
CC bacterium with a recombinant viral vector that encodes a polynucleotide
CC sequence for a mismatch endonuclease, growing the host so that the
CC polynucleotide is expressed, and extracting the mismatch endonuclease
CC enzyme from the host. The method is useful for making mismatch
CC endonuclease enzymes, for obtaining peptides and polynucleotides with
CC desired functional properties and for detecting mutations. The mismatch
CC endonuclease enzymes are useful in gene shuffling technology for

CC developing new genes, in detecting single nucleotide polymorphisms for
CC e.g. detecting evidence of cancer susceptibility, or in redistributing
CC sequence variations between non-identical polynucleotide sequences. The
CC present sequence represents the wild type green fluorescent protein, GFP,
CC gene construct.
XX
SQ Sequence 3637 BP; 959 A; 879 C; 868 G; 931 T; 0 U; 0 Other;
Query Match 72.5%; Score 2656.4; DB 11; Length 3637;
Best Local Similarity 84.0%; Pred. No. 0;
Matches 3079; Conservative 0; Mismatches 556; Indels 29; Gaps 6;
QY 1. CTAAATTTGAGCGTTAAATATTATTTGTTAAATTCGCGTTAAATTTTGTAAATCAGCTC 60
DB 3637 CTAAATTTGAGCGTTAAATATTATTTGTTAAATTCGCGTTAAATTTTGTAAATCAGCTC 3578
QY 61 ATTTTAAACCAATAGGCGAAATCGCAAAATCCCTTATAAATCAAAAGNATAGACCGA 120
DB 3577 ATTTTAAACCAATAGGCGCAAAATCGCAAAATCCCTTATAAATCAAAAGNATAGACCGA 3518
QY 121 GATAGGTTGAGTGTGTTTCCAGTTTGGAAACAGAGTCCACTATTAAAGAACGTGACTC 180
DB 3517 GATAGGTTGAGTGTGTTTCCAGTTTGGAAACAGAGTCCACTATTAAAGAACGTGACTC 3458
QY 181 CAAGTCAAAAGGCGGAAACCCGTCATCAGGGCGATGCGCCCACTACGTGAACCATCACC 240
DB 3457 CAAGTCAAAAGGCGGAAACCCGTCATCAGGGCGATGCGCCCACTACGTGAACCATCACC 3398
QY 241 CTAAATCAAGTTTTTGGGTCGAGGTGCGTAAAGCACTAAATTCGGAACCTTAAGGGAG 300
DB 3397 CTAAATCAAGTTTTTGGGTCGAGGTGCGTAAAGCACTAAATTCGGAACCTTAAGGGAG 3338
QY 301 CCCCCGATTTAGAGCTTGAACGGGAAAGCCGCGAAACGTCGGCGAGAAAAGGGAAGAA 360
DB 3337 CCCCCGATTTAGAGCTTGAACGGGAAAGCCGCGAAACGTCGGCGAGAAAAGGGAAGAA 3278
QY 361 AGCGAAGAGCGGCGCTAGGCGCTGCAAGTGTAGCGGTCACTGCGGTGCGGTGCGGTGCG 420
DB 3277 AGCGAAGAGCGGCGCTAGGCGCTGCAAGTGTAGCGGTCACTGCGGTGCGGTGCGGTGCG 3218
QY 421 CACACCGCGCGCTTAATGCGCGCTACAGGGCGGTCCCATTTGCCCATTCAGGCTGCG 480
DB 3217 CACACCGCGCGCTTAATGCGCGCTACAGGGCGGTCCCATTTGCCCATTCAGGCTGCG 3158
QY 481 CAACTGTTGGGAAGGCGGATCGGTGCGGCTCTTTCGCTATTACGCCAGTGGCGAAAGG 540
DB 3157 CAACTGTTGGGAAGGCGGATCGGTGCGGCTCTTTCGCTATTACGCCAGTGGCGAAAGG 3098
QY 541 GGGATGCTGCAAGGCGATTAAGTTGGGTAAAGCGCAGGGTTTTCCAGTCAGCAGCTTG 600
DB 3097 GGGATGCTGCAAGGCGATTAAGTTGGGTAAAGCGCAGGGTTTTCCAGTCAGCAGCTTG 3038
QY 601 TAAACGACGCGCAGTGAAGCGGCTCGTTTCACTTACGCTTTTGAACCGGTGAGGACCG 660
DB 3037 TAAACGACGCGCAGTGAAGCGGCTCGTTTCACTTACGCTTTTGAACCGGTGAGGACCG 2978
QY 661 GCAGACTCCGCTGCAAAATGTTTTCAGCGCTGATGAGCAGATCAAGATGCTCGACAC 720
DB 2977 CCGCGTGGCGCGCTCTAGAACTAGTGGATCCCGCGGCTGCAGGAATCTTATTGTT 2918
QY 721 GCTGCAGAACACGCGAGCTAGATAACCTCCTAGAAAGATATCATATTGTGACAGTACGTTAA 780
DB 2917 ATAGTTTATCCATGCCATGTTAAATCCAGCAGCTGTTACAAACTCAAGAAAGGACCATGT 2858
QY 781 AGATAATC--ATGCGTAAATTTGACCGCATGGGATCTGTAATACGACTCACTATAGGCGA 838
DB 2857 GGTCTCTCTTTTGTGTTGGGATCTTTTCGAAAGGCGCAGATTGTGTGGACAGGTAATGTTGT 2798
QY 839 ATTGGGTACCGGCGCCCGCTCGAGGTCCAGGTATCGATAGCTTGATATCGAAATTCCT 898
DB 2797 CTGGTAAAGGACAGGCGCCATCGCCAAATGGAGTATTTTGTGATTAATGGTCTGCTAGTT 2738
QY 899 GCAGCCCGGGGATCCACTAGTTTCTAGAGCGCGCCGCCACCGCGGTGGAGCTCCAGCTTTT 958

Db 2737 GAACGCTCCATCTTCAATGTTGTGCTAAATTTGAAAGTTAACTTTGATTCATCTTTT 2678
QY 959 GTTCCCTTTAGTGAGGGTTAAATAGATCCATCGCTCAATTTTACGCAGACTATCTTTCT 1018
Db 2677 GTTGTGCTGCCAAGATGATACATTTGTGGA-GTTATAGTTGATTTCCAAATTTGTCCA 2619
QY 1019 AGGGTTAAATCTAGCTGCATCAGATCATATCGTCGGGTCTTTTTCGGGCTCAGTCATCG 1078
Db 2618 AGAATGTTTCCATCTTCTTAAATCAATACCTTTTAACTCGATTC-----TATTAAANA 2564
QY 1079 CCCAAGCTGGGCTATCTGGGCAATCGGGAGGAAGAACCCGTGCTTTTCCCGGAGGT 1138
Db 2563 GGGTATCACCTTTCAAATTTGACTTC-----AGCACGTGCTTGTAGTTCGCGTC 2515
QY 1139 TGAAGGGGATGGAAAGAGTTTGGCAGGATGACTGCTGCTGCATGAGCTTGAGGGAAA 1198
Db 2514 ATCTTTGAAAAATATAGTTCTTTCTGTATACAACTTCGGGCATGGCACTCTTGAANAAA 2455
QY 1199 AGCGACGTTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCTTTAAACGGTGAACCT 1258
Db 2454 GTCATGCCGTTTCATATGATCTGGGTACTTGAAGAAGCATTTGAAACACCATATAGANAAGT 2395
QY 1259 GTTCGTTCAAGCCACTTGGGATACCAATTCGTGCGGCTTTTCCGGACACAGTTCCGGAT 1318
Db 2394 AGTGACAAGTGTGGCCATGGAAACAGGTAGTTTTCAGTAGTGCAAAATAAAATTTAAGGGT 2335
QY 1319 GGTACAGCGAAGCGATCAGCAACCCGNAACATACCGGCGACAGCCGGAACCTGCGTGC 1378
Db 2334 AAGTTTTC-----CGTATGTTGCATCACCTTCACCTCTCCACTGACAGAAAATTTGTGC 2280
QY 1379 CGGTGTGCAGATTAATGACAGCGGTGCGGCTGGGATATTACGTACAGCGAGCAGGGTA 1438
Db 2279 CCAITTAACATACCAATCTAATTCACAAGAANTTGGACAACTCCAGTGAAGAAGTTCTTCT 2220
QY 1439 TCTCGCTGGATGCGCGAGAAATGGAATGGAATACCCCGTGAGTTTACCGCGGCGCGC 1498
Db 2219 CTTTACTCATCGGTACCCAGCTTTTGTGTTCCCTTTAGTGAGGGTTA-----ATTGCGCGC 2165
QY 1499 TTGGGTAATCATGTCATAGCTGTTTCTGTGTGAATTTGTTATCCGCTCACAATTTCA 1558
Db 2164 TTGGCGTAATCATGTCATAGCTGTTTCTGTGTGAATTTGTTATCCGCTCACAATTTCA 2105
QY 1559 CACAACATACGAGCGGAGACATAAGTGTAAAGCTGGGGTGCCTAATGAGTGAGCTAA 1618
Db 2104 CACAACATACGAGCGGAGACATAAGTGTAAAGCTGGGGTGCCTAATGAGTGAGCTAA 2045
QY 1619 CTCACATTAATTCGCTGCTCATGCGCGCTTTTCCAGTCGGGAAAACCTGTCGTGCCAG 1678
Db 2044 CTCACATTAATTCGCTGCTCATGCGCGCTCACTGCGCGCTTTCCAGTCGGGAAAACCTGTCGTGCCAG 1985
QY 1679 CTGCATTAATGAATCGGCCAACCGCGGGGAGAGGGGTTTCGGTATTTGGGCGCTCTTCC 1738
Db 1984 CTGCATTAATGAATCGGCCAACCGCGGGGAGAGGGGTTTCGGTATTTGGGCGCTCTTCC 1925
QY 1739 GCTTCTCGCTCACTGACTGCTGCGCTCGGTGCTTTGCGCTGCGCGAGCGGTATCAGCT 1798
Db 1924 GCTTCTCGCTCACTGACTGCTGCGCTCGGTGCTTTGCGCTGCGGGAGCGGTATCAGCT 1865
QY 1799 CACTCAAAAGCGGTAATACGGTTATCCACAGAAATCAGGGGATTAACGAGGAAGAACAATG 1858
Db 1864 CACTCAAAAGCGGTAATACGGTTATCCACAGAAATCAGGGGATAACGAGGAAGAACAATG 1805
QY 1859 TGAGCAAAAGGCGAGAAAGGCGAGAACCGTAAAGGCCGCTTGTGCGGCTTTTTC 1918
Db 1804 TGAGCAAAAGGCGAGCAAAAGGCGAGAACCGTAAAGGCCGCTTGTGCGGCTTTTTC 1745
QY 1919 CATAGGCTCCGCCCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGA 1978
Db 1744 CATAGGCTCCGCCCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGA 1685
QY 1979 AACCCGACAGGACTATAAAGATACGAGCGTTTCCCTCTGGAAAGCTTCCTCTCGTGGCTCT 2038

Db 1684 AACCCGACAGGACTATAAAGATACACAGCGTTTCCCCTCGGAAGCTCCCTCTGTCGCTCT 1625
QY 2039 CCTGTTTCGACCCCTCGCTTACCGGATACCTGTCCGCTTTCTCCTCTCGGGAAGCGTG 2098
Db 1624 CTTGTTCCGACCCCTCGCTTACCGGATACCTGTCCGCTTTCTCCTCTCGGGAAGCGTG 1565
QY 2099 GCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCCGGTGTAGGTGCTTCGCTCCAAAG 2158
Db 1564 GCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCCGGTGTAGGTGCTTCGCTCCAAAG 1505
QY 2159 CTGGGCTGTGTGCAGCAAAACCCCGTTCAAGCCGAGCCGCTGCGCTTATTCGGTAACTAT 2218
Db 1504 CTGGGCTGTGTGCAGCAAAACCCCGTTCAAGCCGAGCCGCTGCGCTTATTCGGTAACTAT 1445
QY 2219 CGTCTTGTAGTCAAAACCCCGTTAAGACACACTTATCGCACTGCGCAGCAGCCACTGGTAAC 2278
Db 1444 CGTCTTGTAGTCAAAACCCCGTTAAGACACACTTATCGCACTGCGCAGCAGCCACTGGTAAC 1385
QY 2279 AGGATTAGCAGAGCGAGGTATGTAGGCGGTGTACAGAGTTCTTGAAGTGTGGGCTAAAC 2338
Db 1384 AGGATTAGCAGAGCGAGGTATGTAGGCGGTGTACAGAGTTCTTGAAGTGTGGGCTAAAC 1325
QY 2339 TACGGCTACATAGAAAGACAGATTTTGGTATCTCGGCTCTGCTGAGCCAGCTTACTTTC 2398
Db 1324 TACGGCTACATAGAAAGACAGATTTTGGTATCTCGCTCTGCTGAGCCAGCTTACTTTC 1265
QY 2399 GGAAGAAAGTGTGTAGCTCTTGATCCGCAAAACAAACCCAGCTGGTAGCGGTGGTTTT 2458
Db 1264 GGAAGAAAGTGTGTAGCTCTTGATCCGCGCAAAACAAACCCAGCTGGTAGCGGTGGTTTT 1205
QY 2459 TTTGTTTCAAGCAGCAGATTAACGCGCAAAAAAAGGATCTCAAGAAAGATCCTTTGATC 2518
Db 1204 TTTGTTTCAAGCAGCAGATTAACGCGCAAAAAAAGGATCTCAAGAAAGATCCTTTGATC 1145
QY 2519 TTTTCTACGGGCTGTGACGCTCAGTGGAACGAAACTCAGCTTAAGGAGTTTGGTCATG 2578
Db 1144 TTTTCTACGGGCTGTGACGCTCAGTGGAACGAAACTCAGCTTAAGGAGTTTGGTCATG 1085
QY 2579 AGATTATCAAAAGGATCTTCACCTAGATCTCTTTTAAATTTAAATTTAAATTTAAATCA 2638
Db 1084 AGATTATCAAAAGGATCTTCACCTAGATCTCTTTTAAATTTAAATTTAAATTTAAATCA 1025
QY 2639 ATCTAAAGTATATATGAGTAAACTTTGGTGTGACAGTTACCAATGTCTTAATCAGTGAGGCA 2698
Db 1024 ATCTAAAGTATATATGAGTAAACTTTGGTGTGACAGTTACCAATGTCTTAATCAGTGAGGCA 965
QY 2699 CCTATCTCAGCGATCTGTCTATTTGTTTCATCCATAGTTGCTGACTCCCGCTCGGTAG 2758
Db 964 CCTATCTCAGCGATCTGTCTATTTCTGTTTCATCCATAGTTGCTGACTCCCGCTCGGTAG 905
QY 2759 ATAACTACGATACGGGAGGCTTACCATCTGGCCCCAGTGTGCAATGATACCGGAGAC 2818
Db 904 ATAACTACGATACGGGAGGCTTACCATCTGGCCCCAGTGTGCAATGATACCGGAGAC 845
QY 2819 CCACGCTCAACGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGCCCGAGCGC 2878
Db 844 CCACGCTCAACGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGCCCGAGCGC 785
QY 2879 AGAAGTGTCTCGAACTTTATCCGCTCCATCCAGTCTATTTAATTTGTTCCCGGAGCT 2938
Db 784 AGAAGTGTCTCGAACTTTATCCGCTCCATCCAGTCTATTTAATTTGTTCCCGGAGCT 725
QY 2939 AGAGTAAGTGTCCCGAGTTAATAGTTTGGCAACCTTTGGCCATTTGCTATGCTACAGGCATC 2998
Db 724 AGAGTAAGTGTCCCGAGTTAATAGTTTGGCAACCTTTGGCCATTTGCTATGCTACAGGCATC 665
QY 2999 GTGGTGTACGCTCGTGTGTTGGTATGGCTTCATTCACTCCGCTTCCCAACGATCAAGG 3058
Db 664 GTGGTGTACGCTCGTGTGTTGGTATGGCTTCATTCACTCCGCTTCCCAACGATCAAGG 605
QY 3059 CGAGTTACATGATCCCCCATGTTGTGCAAAAAGCGGTAGTCTCTCGGTCTCTCGGATC 3118
Db 604 CGAGTTACATGATCCCCCATGTTGTGCAAAAAGCGGTAGTCTCTCGGTCTCTCGGATC 545

QY 3119 GTTGTGAGAAAGTGGCGGAGTGTATATCACTCATGTTATGGCAGCACTGCATAAT 3178
Dd |||||
Dd 544 GTTGTGAGAAAGTGGCGGAGTGTATATCACTCATGTTATGGCAGCACTGCATAAT 485
QY 3179 TCTTCTACTGCTCATGTCATCCGTAAGATGCTTTCTGTGACTGCTGAGTACTCAACCAAG 3238
Dd |||||
Dd 484 TCTTCTACTGCTCATGTCATCCGTAAGATGCTTTCTGTGACTGCTGAGTACTCAACCAAG 425
QY 3239 TCATTTCTGAGAAVAGTGTATGCGGCGACCGAGTTGCTTCTGCGCGGCTCAATACCGGAT 3298
Dd |||||
Dd 424 TCATTTCTGAGAAVAGTGTATGCGGCGACCGAGTTGCTTCTGCGCGGCTCAATACCGGAT 365
QY 3299 AATACCGCGCCACATGACGAACTTTAAAGTGCTCATCTTTGGAAAAACGTTCTTCGGGG 3358
Dd |||||
Dd 364 AATACCGCGCCACATGACGAACTTTAAAGTGCTCATCTTTGGAAAAACGTTCTTCGGGG 305
QY 3359 CGAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAAACCCACTCGTGCA 3418
Dd |||||
Dd 304 CGAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAAACCCACTCGTGCA 245
QY 3419 CCCAAGTCTTTCAGCATCTTTTACTTTTCCAGGCGTTTCTGGGTGAGCAAAAACAGGA 3478
Dd |||||
Dd 244 CCCAAGTCTTTCAGCATCTTTTACTTTTCCAGGCGTTTCTGGGTGAGCAAAAACAGGA 195
QY 3479 AGCAAAATGCCCCCAAAAAGGAATAAGGCGACACGGAATGTTGAATCTCATACTC 3538
Dd |||||
Dd 184 AGCAAAATGCCCCCAAAAAGGAATAAGGCGACACGGAATGTTGAATCTCATACTC 125
QY 3539 TTCTCTTTTCAATATTATTTGAAGCATTTATCAGGTTTATGTTCTCATGAGCGATACATA 3598
Dd |||||
Dd 124 TTCTCTTTTCAATATTATTTGAAGCATTTATCAGGTTTATGTTCTCATGAGCGATACATA 65
QY 3599 TTTGAATGTTATTTAGAAAAATAAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAAGTG 3658
Dd |||||
QY 3659 CCAC 3662
Dd |||||
Dd 4 CCAC 1

RESULT 13
ADP26615/C
ID ADP26615 standard; DNA; 3637 BP.
XX
AC ADP26615;
XX
XX
DT 26-AUG-2004 (first entry)
DE
DE Green fluorescent protein (GFP) plasmid DNA #1.
XX
XX Sequence variation; heteroduplex; transcription; DNA integration;
KW ribozyme expression; gene; ds; green fluorescent protein; GFP.
XX
OS Aequorea victoria.
OS Synthetic.
XX
XX US2004110130-A1.
XX
PD 10-JUN-2004.
XX
XX 25-OCT-2002; 2002US-00280913.
XX
XX 02-FEB-2001; 2001US-0266386P.
PR 14-FEB-2001; 2001US-0268785P.
PR 01-FEB-2002; 2002US-00066390.
PR 08-AUG-2002; 2002US-0402342P.
XX
XX (LARG-) LARGE SCALE BIOLOGY CORP.
XX
XX Padgett HS, Lindbo JA, Fitzmaurice WP;
XX

DR WPI: 2004-440326/41.
XX Redistributing sequence variations between non-identical polynucleotide
PT sequences, useful for generating improved polynucleotide having a desired
PT characteristic, comprises making a heteroduplex and introducing a nick.
XX
PS Example 5; SEQ ID NO 31; 75bp; English.
XX
CC The invention relates to an in vitro method of redistributing sequence
CC variations between non-identical polynucleotide sequences, comprising
CC making a heteroduplex polynucleotide from two non-identical
CC polynucleotides, introducing a nick in the second strand at or near a
CC base pair mismatch site, removing the mismatched base(s) from the
CC mismatch site where the nick occurred and using the first strand as a
CC template to replace the removed base(s) with bases that complement the
CC base(s) in the first strand. The invention also relates to an in vitro
CC method of making a population of sequence variants from a heteroduplex
CC polynucleotide sequence, obtaining a polynucleotide sequence encoding a
CC desired functional property and identifying a reassorted DNA molecule
CC encoding a protein with a desired functional property. The method is
CC useful for generating an improved polynucleotide sequence or a population
CC of improved polynucleotide sequences possessing at least one desired
CC phenotypic characteristic (e.g., promotes transcription of linked
CC polynucleotides), where such polynucleotides are useful for expression
CC from a plant, animal, fungal, yeast, or bacterial expression vector, for
CC integration to form a transgenic plant, animal or microorganism, and for
CC expression of a ribozyme. This sequence represents DNA used in the scope
CC of the invention.
XX
SQ Sequence 3637 BP; 959 A; 879 C; 868 G; 931 T; 0 U; 0 Other;
Query Match 72.5%; Score 2656.4; DB 12; Length 3637;
Best Local Similarity 84.0%; Pred. No. 0;
Matches 3079; Conservative 0; Mismatches 556; Indels 29; Gaps 6;
QY 1 CTAAATTTGTAAGCGTTAAATATTTTGTAAATTCGCGTTAAATTTTGTAAATTCAGCTC 60
Dd |||||
Dd 3637 CTAAATTTGTAAGCGTTAAATATTTTGTAAATTCGCGTTAAATTTTGTAAATTCAGCTC 3578
QY 61 ATTTTAAACCAATAGCGCGAAATCGGCAAAATCCCTTATAATCAAAAGAAATAGACCGA 120
Dd |||||
Dd 3577 ATTTTAAACCAATAGCGCGAAATCGGCAAAATCCCTTATAATCAAAAGAAATAGACCGA 3518
QY 121 GATAGGTTGAGTGTGTTTCCAGTTTGGAAACAAGAGTCCACTATTAATAAGAAACCTGACTC 180
Dd |||||
Dd 3517 GATAGGTTGAGTGTGTTTCCAGTTTGGAAACAAGAGTCCACTATTAATAAGAAACCTGACTC 3458
QY 181 CAACGTCAAAGGGCGGAAACCGCTCTATCAGGGCGATGCGCCACTACGTGAACCATCACC 240
Dd |||||
Dd 3457 CAACGTCAAAGGGCGGAAACCGCTCTATCAGGGCGATGCGCCACTACGTGAACCATCACC 3398
QY 241 CTAATCAAGTTTTTTTGGGGTCGAGGTGCGGTAAAGCACTAAATCGGAACCTTAAGGGAG 300
Dd |||||
Dd 3397 CTAATCAAGTTTTTTTGGGGTCGAGGTGCGGTAAAGCACTAAATCGGAACCTTAAGGGAG 3338
QY 301 CCCCCGATTTAGAGCTTGAACGGGGAAAGCCGGCGAAACGTCGCGAGAAAGGAGGGAAGAA 360
Dd |||||
Dd 3337 CCCCCGATTTAGAGCTTGAACGGGGAAAGCCGGCGAAACGTCGCGAGAAAGGAGGGAAGAA 3278
QY 361 AGCGAAGAGGCGGCGCTAGGGCGCTGAGCGTGCAGAGTGTAGCGGTCAAGCTGCGGTACCCAC 420
Dd |||||
Dd 3277 AGCGAAGAGGCGGCGCTAGGGCGCTGAGCGTGCAGAGTGTAGCGGTCAAGCTGCGGTACCCAC 3218
QY 421 CACACCGCGCGCTTTAATGCGCGCTACAGGGCGGTCCCATTCGCCATTCAGGCTGCG 480
Dd |||||
Dd 3217 CACACCGCGCGCTTTAATGCGCGCTACAGGGCGGTCCCATTCGCCATTCAGGCTGCG 3158
QY 481 CAACTGTTGGGAAGGGCGATCGGTGCGGCGCTTCTCGCTATTACGCCAGCTGCGGAAGG 540
Dd |||||
Dd 3157 CAACTGTTGGGAAGGGCGATCGGTGCGGCGCTTCTCGCTATTACGCCAGCTGCGGAAGG 3098
QY 541 GGGATGCTGTCAGAGCGGATTAAGTTGGGTAAAGCCAGCGGTTTTTCCAGTCAGAGCTTG 600
Dd |||||

Dd	3097	GGGATGTGCTGCAAGGCGAATTAAGTTGGGTAAAGCCAGGCGTTTCCAGTCAAGCGTTG	3038
Qy	601	TAAACGACGGCCAGTACGCGCGCTCGTTCAATCAGCTTTTGAACCCGTCGAGACGG	660
Dd	3037	TAAACGACGGCCAGTACGCGCGGTAAACGACTCACTATAGGGGAATTTGAGCTCCA	2978
Qy	661	GCAGACTCCGGTGCAGAAATGTGTTTACAGCGTGTAGAGAGATGAAGATGCTCGACAC	720
Dd	2977	CCGGGTGGCGCGCTCTAGAACTAGTGGATCCCGCGGCTGCAGGAATTTCTATTGT	2918
Qy	721	GCTGCAGAACACGACGTAGATAAACCCTAGAAAGATAATATATGTGACGTACGTTAA	780
Dd	2917	ATAGTTTCAATCCATGCCATGTGTAATCCACAGAGCTGTTACAAACTCAAGAGGACCATGT	2858
Qy	781	AGATAATC--ATGCGTAAATTTGACGCAATGGGATCTGTAAATACGACTCACTATAGGGCA	838
Dd	2857	GGTCTCTCTTTTGGTGGGATCTTTGAAAGGCGAGATTTGTGTGACAGGTAATGGTGT	2798
Qy	839	ATTGGGTACCGGGCCCCCTCGAGGTCCAGCGTATCGATAAGCTTTGATATCGAAATTCCT	898
Dd	2797	CTGGTAAAGGACAGGGCCATCGCAATTTGGAGTATTTTGTGATAATGGTCTGTAGTT	2738
Qy	899	GCAGCCGGGGGATCCACTAGTTCTTAGAGCGCGCCACCGCGTGGAGCTCCAGCTTTT	958
Dd	2737	GAAAGCTTCCATCTTCAATGTGTGCTAAATTTTGAAGTTAACTTTGATTTCCATTTCTTTT	2678
Qy	959	GTTCCTTTTAGTGGGTTAATTAGATCCATGCGTCAATTTTACGACACTATCTTTCT	1018
Dd	2677	GTTTGTCTGCCAAGATATACATTTGTGTGA--GTATAGTTGTATTTCCAAATTTGTGTCCA	2619
Qy	1019	AGGGTTAATCTAGCTGCATCAGGATCATATCGCGGTCTTTTTCGGGCTCAGTCATCG	1078
Dd	2618	AGAAATGTTTCCATCTCTTTAAATCAATACCTTTTAACTCGATTC--TATTAAACAA	2564
Qy	1079	CCCAAGCTGGCGTATCTGGGCATTCGGGGAGGAAGAGCCCGTGTCTTTCCCGCGAGGT	1138
Dd	2563	GGGTATCACTTCAAACCTTGAATTC-----AGCACGTGTCTTTAGTTCCCGTC	2515
Qy	1139	TGAAGCGGATGGAAGAGTTTGGCGAGGATCACTGCTGCTCAATTTGAGTTGAGCGAAA	1198
Dd	2514	ATCCTTTGAAAATATAGTTCTTTCTGTATACATAACCTTCGGGCATGGCACTTTGAAANA	2455
Qy	1199	ACGCACGTTTACCATGATGATTTGGGAAAGTGTGGCCATGCACGCGCTTTTAAACGTTGA	1258
Dd	2454	GTCAATGCCGTTTCAATATGATCTGGGTATCTTGAAGCAATTAACACCATAGAGAAAT	2395
Qy	1259	GTTCGTTACGGCCACCTGGGATACAGTTTCGTGGGCTTTTCCGGACACAGTTCCGGAT	1318
Dd	2394	AGTGACAAGTGTGGCCATGGAAACAGGTAGTTTCCAGTAGTGCATAAATAAATTTAAGGT	2335
Qy	1319	GGTCAGCCGAGCGCATCAGCAACCCGAAATATACCGCGGACAGCCGAACTGCCGTGC	1378
Dd	2334	AAGTTTTC-----CGTATGTTGCATCACCTTCACCCCTCTCCACTGACAGAAATTTGTGC	2280
Qy	1379	CGGTGTGCAGATTAATAGACGCGTGGCGCTGGGATATTACGTACAGCGAGGACGGGTA	1438
Dd	2279	CCATTAAACATCACCATCTAATTTCAACAGAAATTTGGGACAACTCCAGTGAAGTTCTTCT	2220
Qy	1439	TCCTGGCTGGATGCCGCGAGAAATGGACATGGATACCCCGTAGTTTACCCGGCGGGCGGC	1498
Dd	2219	CCTTTACTCATCGTACCCAGCTTTTGTTCCTTTTGTAGTGGGGTTA-----ATTGCGGC	2165
Qy	1499	TTGGCGTAAATCATGTTCTAGCTGTTTCTCTGTGTGAATTTGTTATCCGCTCACAATTTCA	1558
Dd	2164	TTGGCGTAAATCATGTTCTAGCTGTTTCTGTGTGAATTTGTTATCCGCTCACAATTTCA	2105
Qy	1559	CACACATACGACCGGAGCATAAAGTGTAAAGCTTGGGGTGCCTTAATGAGTGAGCTAA	1618
Dd	2104	CACACATACGACCGGAGCATAAAGTGTAAAGCTTGGGGTGCCTTAATGAGTGAGCTAA	2045
Qy	1619	CTCACATTAATTTGGTGGCTCACTGCGCGCTTTCCAGTCGGGAAAACCTGTCGTGCCAG	1678
Dd	2044	CTCACATTAATTTGGTGGCTCACTGCGCGCTTTCCAGTCGGGAAAACCTGTCGTGCCAG	1985
Qy	1679	CTGCATTAATTAATCGGCCAACCGCGGGAGAGAGCGGTTTTCGTTATTTGGGCGCTCTTCC	1738
Dd	1984	CTGCATTAATTAATCGGCCAACCGCGGGAGAGAGCGGTTTTCGTTATTTGGGCGCTCTTCC	1925
Qy	1739	GCTTCTCGCTCACTGACTCGCTCGCTCGGTCGCTGCGCTGCGCGAGCGGTATCAGCT	1798
Dd	1924	GCTTCTCGCTCACTGACTCGCTCGCTCGGTCGCTGCGCTGCGCGAGCGGTATCAGCT	1865
Qy	1799	CACTAAAGGCGGTAAATACGGTTATCCACAGNATCAGGGGATAACGCGAGGAAGACATG	1858
Dd	1864	CACTAAAGGCGGTAAATACGGTTATCCACAGNATCAGGGGATAACGCGAGGAAGACATG	1805
Qy	1859	TCAGCAAAAGGCGAGCAAAAGGCGAGGAACCTGAAAAAGGCGCGTTCGTCGCGCTTTTC	1918
Dd	1804	TCAGCAAAAGGCGAGCAAAAGGCGAGGAACCTGAAAAAGGCGCGTTCGTCGCGCTTTTC	1745
Qy	1919	CATAGGCTCCGCCCCCTTGAAGAGATCACAAAAAATCGACGCTCAAGTCAGAGGTGGCGA	1978
Dd	1744	CATAGGCTCCGCCCCCTTGAAGAGATCACAAAAAATCGACGCTCAAGTCAGAGGTGGCGA	1685
Qy	1979	RACCCGACAGGACTATAAGATACAGGGGTTTCCCGCTGGAGCTCCCTCGTGGCTCT	2038
Dd	1684	AAACCCGACAGGACTATAAGATACAGGGGTTTCCCGCTGGAGCTCCCTCGTGGCTCT	1625
Qy	2039	CCTGTTCCGACCCCTTACCGGATACTGTCCGCTTCTTCTCCCTTCGGGAAAGCGTG	2098
Dd	1624	CCTGTTCCGACCCCTTACCGGATACTGTCCGCTTCTTCTCCCTTCGGGAAAGCGTG	1565
Qy	2099	GGCTTTCTCATPATGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTTCGCTCCAAG	2158
Dd	1564	GGCTTTCTCATPATGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTTCGCTCCAAG	1505
Qy	2159	CTGGGCTGTGACAGAACCCCGTTACGCCGACCGCTGCGCTTATCCCGTAACCTAT	2218
Dd	1504	CTGGGCTGTGACAGAACCCCGTTACGCCGACCGCTGCGCTTATCCCGTAACCTAT	1445
Qy	2219	CGTCTTGAGTCAAAACCCCGTTAAGACACGACTTATCGCACCTGGCAGCAGCCACTGGTAAC	2278
Dd	1444	CGTCTTGAGTCAAAACCCCGTTAAGACACGACTTATCGCACCTGGCAGCAGCCACTGGTAAC	1385
Qy	2279	AGGATTAGCAGACGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGTGGCTTAAC	2338
Dd	1384	AGGATTAGCAGACGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGTGGCTTAAC	1325
Qy	2339	TACGGCTACCTAGAGGACAGTATTTGCTATCTCGCTCTGCTGACGACAGTACCTTC	2398
Dd	1324	TACGGCTACCTAGAGGACAGTATTTGCTATCTCGCTCTGCTGACGACAGTACCTTC	1265
Qy	2399	GGAAAAAGATTGGTAGTCTTTGATCCGCGCAAAACAAACCCCGCTGGTAGCGGTGTTTT	2458
Dd	1264	GGAAAAAGATTGGTAGTCTTTGATCCGCGCAAAACAAACCCCGCTGGTAGCGGTGTTTT	1205
Qy	2459	TTTGTTCGACAGCAGCATTAACGCGCAAAAGGATCTCAAGAGATCTTTTGATC	2518
Dd	1204	TTTGTTCGACAGCAGCATTAACGCGCAAAAGGATCTCAAGAGATCTTTTGATC	1145
Qy	2519	TTTTCTACGGGCTCTGACGCTCAGTGGAAACGAAACTCACGTTAAGGGATTTTGGTCATG	2578
Dd	1144	TTTTCTACGGGCTCTGACGCTCAGTGGAAACGAAACTCACGTTAAGGGATTTTGGTCATG	1085
Qy	2579	AGATTATCAAAAGGATCTTTCACCTAGATCTTTTAAATTTAAATTTAAATTTAAATCA	2638
Dd	1084	AGATTATCAAAAGGATCTTTCACCTAGATCTTTTAAATTTAAATTTAAATTTAAATCA	1025
Qy	2639	ATCTAAAGTATATATGAGTAAACTTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCA	2698
Dd	1024	ATCTAAAGTATATATGAGTAAACTTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCA	965
Qy	2699	CCTATCTCAGCGATCTGTCTATTTCTTCATCTAGTTGCTGACTCCCGCTCGTGTAG	2758
Dd	964	CCTATCTCAGCGATCTGTCTATTTCTTCATCTAGTTGCTGACTCCCGCTCGTGTAG	905

QY	2759	ATRAACTACGATACGGAGGGCTTACCATCTGGCCCCAGTGTCTGCAATGATACCGCAGAC	2818
Db	904	ATAAATACGATACGGAGGGCTTACCATCTGGCCCCAGTGTCTGCAATGATACCGCAGAC	845
QY	2819	CCACGCTCACGGCTCCAGATTATCAGCAATAAACACGACGCGGAAAGGCGCAGCGC	2878
Db	844	CCACGCTCACGGCTCCAGATTATCAGCAATAAACACGACGCGGAAAGGCGCAGCGC	785
QY	2879	AGAAGTGGTCTCTGCAACTTTATCCGGCTCCATCCAGTCTATTAATTTGTTGCCGGGAAGCT	2938
Db	784	AGAAGTGGTCTCTGCAACTTTATCCGGCTCCATCCAGTCTATTAATTTGTTGCCGGGAAGCT	725
QY	2939	AGAGTAAGTAGTTGCGCAGTTAATAGTTTGGCGCAAGCTTGTGCCAATGCTACAGGCATC	2998
Db	724	AGAGTAAGTAGTTGCGCAGTTAATAGTTTGGCGCAAGCTTGTGCCAATGCTACAGGCATC	665
QY	2999	GTGGTGTACGCTCGTCGTTGGTATGGCTTCAATTCAGCTTCGGTTCCCAACGATCAAGG	3058
Db	664	GTGGTGTACGCTCGTCGTTGGTATGGCTTCAATTCAGCTTCGGTTCCCAACGATCAAGG	605
QY	3059	CGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCTCTCGGTCTCTCGGATC	3118
Db	604	CGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCTCTCGGTCTCTCGGATC	545
QY	3119	GTGTGAGAAAGTAAGTTGGCCGAGTGTATCACTCATGTTATGCGCAGCACTGCATAT	3178
Db	544	GTGTGAGAAAGTAAGTTGGCCGAGTGTATCACTCATGTTATGCGCAGCACTGCATAT	485
QY	3179	TCCTTTACTGTCAATGCTCCGTAAGTCTTTTCTGACTGCTGCTGCTCAACCAAG	3238
Db	484	TCCTTTACTGTCAATGCTCCGTAAGTCTTTTCTGACTGCTGCTGCTCAACCAAG	425
QY	3239	TCATTTCTGAGAAATGATGCGGCGACCGAGTTGCTCTTTCGCCGCGTCAATACCGGAT	3298
Db	424	TCATTTCTGAGAAATGATGCGGCGACCGAGTTGCTCTTTCGCCGCGTCAATACCGGAT	365
QY	3299	AATACCGCGCACATAGCAGAACTTTAAAGTGCTCATCATTTGGAAAAAGTCTTCGGGG	3358
Db	364	AATACCGCGCACATAGCAGAACTTTAAAGTGCTCATCATTTGGAAAAAGTCTTCGGGG	305
QY	3359	CGAAACTCTCAGGATCTTACCGCTGTTGAGATCCAGTTCGATTAACCCACTCGTGA	3418
Db	304	CGAAACTCTCAGGATCTTACCGCTGTTGAGATCCAGTTCGATTAACCCACTCGTGA	245
QY	3419	CCCACTGATCTTCAGCATCTTTTACTTTTACCAGCGGTTTCTGGGTGAGCAAAAAACGGA	3478
Db	244	CCCACTGATCTTCAGCATCTTTTACTTTTACCAGCGGTTTCTGGGTGAGCAAAAAACGGA	185
QY	3479	AGCAAAATGCCGCAAAAAAGGAAATAAGGCGACACCGGAAATGTTGAAATCTCATCTC	3538
Db	184	AGCAAAATGCCGCAAAAAAGGAAATAAGGCGACACCGGAAATGTTGAAATCTCATCTC	125
QY	3539	TTCCCTTTTCAATATTATTAAGCATTTATCAGGTTATGTTCTCATGACCGATACATA	3598
Db	124	TTCCCTTTTCAATATTATTAAGCATTTATCAGGTTATGTTCTCATGACCGATACATA	65
QY	3599	TTTGAATGTATTAGAAAAATAAAATAAGGAGTTCCGCGCACATTTCCCGAAAAAGTG	3658
Db	64	TTTGAATGTATTAGAAAAATAAAATAAGGAGTTCCGCGCACATTTCCCGAAAAAGTG	5
QY	3659	CCAC 3662	
Db	4	CCAC 1	

RESULT 14
ADQ88437/C
ID ADQ88437 standard; DNA; 3637 BP.
XX
AC ADQ88437;
XX
DT 18-NOV-2004 (first entry)
XX

DE	pbSWTGFPP DNA encoding Jellyfish wild-type GFP protein.	
XX		
KW	Molecular biology; jellyfish; green fluorescent protein; GFP; ds.	
XX		
OS	Aequorea victoria.	
XX		
PN	US2004142433-A1.	
XX		
PD	22-JUL-2004.	
XX		
PF	10-OCT-2003; 2003US-00684134.	
XX		
PR	02-FEB-2001; 2001US-0266386P.	
PR	14-FEB-2001; 2001US-0268785P.	
PR	01-FEB-2002; 2002US-00066390.	
PR	08-AUG-2002; 2002US-0402342P.	
PR	21-AUG-2002; 2002US-00226372.	
PR	25-OCT-2002; 2002US-00280913.	
PR	08-AUG-2003; 2003US-00637758.	
XX		
PA	(PADG/) PADGETT H S.	
PA	(FITZ/) FITZMAURICE W P.	
PA	(LIND/) LINDBO J A.	
PA	(VAEW/) VAETHONGS A A.	
PA	(VOJD/) VOJDANI F S.	
PA	(SMIT/) SMITH M L.	
XX		
PI	Padgett HS, Fitzmaurice WP, Lindbo JA, Vaethongs AA, Vojdani FS;	
PI	Smith ML;	
XX		
XX	WPI; 2004-552565/53.	
DR		
XX		
PT	Preparing variant polynucleotides having different nucleotide sequences	
PT	from at least two parent polynucleotides, useful in molecular biology, in	
PT	particular for generating populations of related nucleic acid molecules.	
XX		
PS	Example 5; SEQ ID NO 31; 81pp; English.	
XX		
CC	The invention relates to a method of preparing a variant polynucleotide	
CC	having a different nucleotide sequence from at least two parent	
CC	polynucleotides. The method involves preparing at least one heteroduplex	
CC	between the two parent polynucleotides, cleaving at least one	
CC	polynucleotide strand in the heteroduplex at a mismatch site to form a	
CC	cleavage site, replacing at least one nucleotide on at least one strand	
CC	at or near the cleavage site where at least one of the strands has a	
CC	different nucleotide sequence from either of the at least two parent	
CC	polynucleotides. Methods and compositions of the invention are useful in	
CC	the field of molecular biology in particular for generating populations	
CC	of related nucleic acid molecules. The present sequence is pbSWTGFPP	
CC	plasmid DNA encoding Jellyfish wild-type GFP (green fluorescent protein).	
CC	This sequence is used in the exemplification of the invention	
XX		
SQ	Sequence 3637 BP; 959 A; 879 C; 868 G; 931 T; 0 U; 0 Other;	
	Query Match 72.5%; Score 2656.4; DB 13; Length 3637;	
	Best Local Similarity 84.0%; Pred. No. 0;	
	Matches 3079; Conservative 0; Mismatches 556; Indels 29; Gaps 6;	
QY	1 CTAAATTGTAAGCGTTAATATTGTTGTTAAATTCGCGTTAAATTTGTTGTTAAATTCAGTCT	60
Db	3637 CTAAATTGTAAGCGTTAATATTGTTGTTAAATTCGCGTTAAATTTGTTGTTAAATTCAGTCT	3578
QY	61 ATTTTAAACCAATAGCGCAAAATCGCAAAATCCCTTATAATCAAAAGAAATAGACCGA	120
Db	3577 ATTTTAAACCAATAGCGCAAAATCGCAAAATCCCTTATAATCAAAAGAAATAGACCGA	3518
QY	121 GATAGGTTGAGTGTGTTCCAGTTTGGAAACAGAGTCCACTATTAAAGAACTGAGTCT	180
Db	3517 GATAGGTTGAGTGTGTTCCAGTTTGGAAACAGAGTCCACTATTAAAGAACTGAGTCT	3458
QY	181 CAACGTCAAAGGGCGGAAACCGTCTATCAGGGCGATGCCCCACTACGTGAACCATCACC	240
Db	3457 CAACGTCAAAGGGCGGAAACCGTCTATCAGGGCGATGCCCCACTACGTGAACCATCACC	3398

QY 241 CTAATCAAGTTTTTTTGGGTCGAGTGCCTGTAAGCACTAAATCGGAACCCCTAAAGGGAG 300
DB |||||
3397 CTAATCAAGTTTTTTGGGGTCGAGTGCCTGTAAGCACTAAATCGGAACCCCTAAAGGGAG 3338
QY 301 CCCCCGATTTAGAGCTTTGACGGGGAAGCCGGGGAACGTTGGCGAGAAAGGAAGGAAGAA 360
DB |||||
3337 CCCCCGATTTAGAGCTTTGACGGGGAAGCCGGGGAACGTTGGCGAGAAAGGAAGGAAGAA 3278
QY 361 AGCGAAGCAGCGGGCGCTAGGGCGCTGCAAGTGTAGCGGTCAAGCTCGCGGTAAACCA 420
DB |||||
3277 AGCGAAGCAGCGGGCGCTAGGGCGCTGCAAGTGTAGCGGTCAAGCTCGCGGTAAACCA 3218
QY 421 CACACCCGCGCGCTTAATGCGCGCTACAGGGCGCTCCCATTCGCCATTTCAGGCTGCG 480
DB |||||
3217 CACACCCGCGCGCTTAATGCGCGCTACAGGGCGCTCCCATTCGCCATTTCAGGCTGCG 3158
QY 481 CAACTGTTGGGAAGGGCGATCGGTGCGGCCCTCTTCGCTATTACGCACTCGCGGAAGG 540
DB |||||
3157 CAACTGTTGGGAAGGGCGATCGGTGCGGCCCTCTTCGCTATTACGCACTCGCGGAAGG 3098
QY 541 GGGATGCTGCAAGCGCAATTAAGTTGGGTAAAGCGGCTTTCCAGTCAAGCACTG 600
DB |||||
3097 GGGATGCTGCAAGCGCAATTAAGTTGGGTAAAGCGGCTTTCCAGTCAAGCACTG 3038
QY 601 TAAAAACGACGGCGAGTACGGCGCTCGTTCAATTCACGTTTTTGAACCCCGTGGAGACGG 660
DB |||||
3037 TAAAAACGACGGCGAGTACGGCGCTCGTTCAATTCACGTTTTTGAACCCCGTGGAGTCCA 2978
QY 661 GCAGACTCGCGTGCAAAATGTTTTTACAGCTGATGGAGCAGATGAAGACTCTCGACAC 720
DB |||||
2977 CCGCGGTGCGCGCGCTCTAGAACTAGTGGATCCCGCGGCTGCAGGAATCTTATTGTG 2918
QY 721 GCTGCAGAACCGCAGCTAGATTAAACCTAGAAAGATATCATATTGTGACGTACGTTAA 780
DB |||||
2917 ATAGTTTCAATCGCATGCTGTAATCCCGAGCGTGTTCACAACTCAAGAGGACCATGT 2858
QY 781 AGATAATC--ATGCGGTAAATTAAGCATGCGATCTGTAAATACGACTCACTATAGGGCA 838
DB |||||
2857 GGTCTCTCTTTTCTGTTGGGATCTTTTCGAAAGGCGAGATTGTGTGACAGGTAATGGTGT 2798
QY 839 ATTGGGTACCGGGCCCCCTCGAGGTGCA CGGTATCGATAAGCTTGATATCGAATTCCT 898
DB |||||
2797 CTGGTAAAGGACAGGGGCCATCGCCAAATGGAGTATTTTGTGATAAATGGTCTGTAGTT 2738
QY 899 GCAGCCGGGGGATCCACTAGTTCTAGAGCGGCCCAACCGCGGTGGAGCTCCAGCTTTT 958
DB |||||
2737 GAACGCTTCCATCTTCAATGTTGTGTTAAATTTTGAAGTTAACTTTGATTCATTTCTTTT 2678
QY 959 GTTCCCTTTTGTGAGGGTTAAATTAGATCCATGCGTCAATTTTACGCACTATCTTTCT 1018
DB |||||
2677 GTTTGCTGCCATGATGATATACATTTGTGTGA--GTTATAGTTGATTTCCAAATTTGTGTCA 2619
QY 1019 AGGGTTAATCTAGTGCATACAGGATCATATCTCGGGTCTTTTTCGGGCTCAGTCACTG 1078
DB |||||
2618 AGAATGTTTCCATCTCTTTAAATCAATACCTTTTAACTCGAATTC-----TATTAAACAA 2564
QY 1079 CCCAAGCTGGGCTATCTGGGATCGGGGAGGAAGCCGCTGCTTTTCCCGGAGGT 1138
DB |||||
2563 GGGTATACCTTTCAACTTGACTTC-----AGCACGTGCTTGTGATTCCTCCGTC 2515
QY 1139 TGAAGCGGCATGGAAGAGTTTTCGCGAGGATGACTGCTGCTCATTTGACGTTTGAAGCGAAA 1198
DB |||||
2514 ATCTTTGAAAATATAGTTCTTTCTGTGTACATAACCTTCGGGCATGGCACTCTTGAAGAA 2455
QY 1199 AGCGAGTTTACATGATGATTCGGGAAGGTTGGCGATGCAAGCGCTTTTAAACGTTGAAC 1258
DB |||||
2454 GTCATGCGGTTTTCATATGATCTGGGTATCTTTGAAAAGCATTTGAACACCATAGAGAAAGT 2395
QY 1259 GTTTCGTTCAGGCCACCTGGGATACCAATTCGTCGGGCTTTTCCGGACACAGTTCGGAT 1318
DB |||||
2394 AGTGACAAAGTTGGCCATGGAAACAGGTAGTTTTCAGTGTAGTGAATAAATTTAAGGGT 2335

QY 1319 GGTACGCCGAGCGCATCAGCAACCCGAACTATACCGCGACAGACCGGAACCTCGCGTGC 1378
DB |||||
2334 AAGTTTTC-----CGTATGTTGCATCACCTTCACTTCCACTGCACAGAAAATTTGTGC 2280
QY 1379 CGGTGTGCAGATTAATGACAGCGGTGCGCGCTGGGATATTACGTCAGCGAGACGGGTA 1438
DB |||||
2279 CCATTAACATCACCATTAATTAACAAGAATTTGGGCAACTCCAGTGAAGAAAGTTCTTCT 2220
QY 1439 TCTCGCTGGATGCGCGAGAAATGACATGGATACCCGTGAGTTTACCCGGCGGGCGGC 1498
DB |||||
2219 CTTTACTCATCGGTACCCAGCTTTTGTTCCTTTAGTGAGGGTTA-----ATTGGCGGC 2165
QY 1499 TTGGCGTAAATCATGTCTAGTGTCTCTGTGTGAAATTTGTTATCCGCTCAAAATCCA 1558
DB |||||
2164 TTGGCGTAAATCATGTCTAGTGTCTCTGTGTGAAATTTGTTATCCGCTCAAAATCCA 2105
QY 1559 CACACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGTGCTTAATGATGAGCTAA 1618
DB |||||
2104 CACACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGTGCTTAATGATGAGCTAA 2045
QY 1619 CTCACATTAATTTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCTGCCAG 1678
DB |||||
2044 CTCACATTAATTTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCTGCCAG 1985
QY 1679 CTGCATTAATGAAATCGGCCAACCGCGGGGAGAGCGGTTTCGCTATTGGGCGCTCTTCC 1738
DB |||||
1984 CTGCATTAATGAAATCGGCCAACCGCGGGGAGAGCGGTTTCGCTATTGGGCGCTCTTCC 1925
QY 1739 GCTTCTCGTCACTGACTCGCTGCGCTCGGTGCTGCTGCGGGAGCGGTATTCAGCT 1798
DB |||||
1924 GCTTCTCGTCACTGACTCGCTGCGCTCGGTGCTGCTGCGGGAGCGGTATTCAGCT 1865
QY 1799 CACTCAAAAGGGCGTAAATACGGTTATCCACAGATCAGGGGATAACGAGGAAGAACATG 1858
DB |||||
1864 CACTCAAAAGGGCGTAAATACGGTTATCCACAGATCAGGGGATAACGAGGAAGAACATG 1805
QY 1859 TGAGCAAAAGGGCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCTGCTGCGCTTTTTC 1918
DB |||||
1804 TGAGCAAAAGGGCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCTGCTGCGCTTTTTC 1745
QY 1919 CATAGCTTCGCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCA 1978
DB |||||
1744 CATAGCTTCGCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCA 1685
QY 1979 AACCCGACAGGACTATAAGATACCAAGCGTTTCCCTCGAAGCTCCCTCGTGGCTCT 2038
DB |||||
1684 AACCCGACAGGACTATAAGATACCAAGCGTTTCCCTCGAAGCTCCCTCGTGGCTCT 1625
QY 2039 CTTGTTCCGACCCCTGCGCTTACCGGATACCTGTCCGCTTTTCTCCCTTCGGGAAGCGTG 2098
DB |||||
1624 CTTGTTCCGACCCCTGCGCTTACCGGATACCTGTCCGCTTTTCTCCCTTCGGGAAGCGTG 1565
QY 2099 GGGCTTTCTCATGCTCAGCTGTAGGTATCTCAGTTCCGTTCCGTTAGGTGCTTCGCTCCAG 2158
DB |||||
1564 GGGCTTTCTCATGCTCAGCTGTAGGTATCTCAGTTCCGTTCCGTTAGGTGCTTCGCTCCAG 1505
QY 2159 CTGGGCTGTGTGACAGAAACCCCGTTCAAGCCGACCGCTGCGCTTATTCGGTAACTAT 2218
DB |||||
1504 CTGGGCTGTGTGACAGAAACCCCGTTCAAGCCGACCGCTGCGCTTATTCGGTAACTAT 1445
QY 2219 CGTCTTGAGTCCAAACCGGTAAAGACAGACTTATCGCACTGGCAGACCACTGTTAAC 1385
DB |||||
2279 AGGATTACGAGCGAGGTATGTAGGCGTCTACAGAGTTCTTCAAGTGTGGGCTAAC 2338
QY 1384 AGGATTACGAGCGAGGTATGTAGGCGTCTACAGAGTTCTTCAAGTGTGGGCTAAC 1325
DB |||||
2339 TACGGCTACCTAGAAAGCAGTATTTGGTATCTCGGCTCTGCTCAAGCCAGCTTACCTTC 2398
QY 1324 TACGGCTACCTAGAAAGCAGTATTTGGTATCTCGGCTCTGCTCAAGCCAGCTTACCTTC 1265
DB |||||
2399 GGAAGAAAGAGTTGGTGTAGTCTTTGATTCGGGCAAAACAAACCGCTGGTAGCGGTGTTTT 2458

Db 1264 GGAAGAGAGTGGTGTAGCTCTTGATCCGCAAAACAAACACACCGCTGGTAGCGTGGTTTT 1205
QY 2459 TTGTTTGCAGCAGCAGATTAACCGCAGAGAAAAAGAGTCTCAAGAAAGATCTTTGATC 2518
Db 1204 TTGTTTGCAGCAGCAGATTAACCGCAGAGAAAAAGAGTCTCAAGAAAGATCTTTGATC 1145
QY 2519 TTTTCTACGGGGTCTGACGCTCAGTGGAAACGAAACCTCAGTTAAGGGATTTTGGTCATG 2578
Db 1144 TTTTCTACGGGGTCTGACGCTCAGTGGAAACGAAACCTCAGTTAAGGGATTTTGGTCATG 1085
QY 2579 AGATTATCAAAAGAGATCTTCACTAGATCTTTTAAATTAATAATGAAGTTTAAATCA 2638
Db 1084 AGATTATCAAAAGAGATCTTCACTAGATCTTTTAAATTAATAATGAAGTTTAAATCA 1025
QY 2639 ATCTAAAGTATATATGAGTAACCTGGTCTGACAGTTTACCAATGCTTAATCAGTAGAGCA 2698
Db 1024 ATCTAAAGTATATATGAGTAACCTGGTCTGACAGTTTACCAATGCTTAATCAGTAGAGCA 965
QY 2699 CCTATCTACGGATCTGTCTATTTCGTTTCAATCAGTGTGCTTCCCGTGGTAG 2758
Db 964 CCTATCTACGGATCTGTCTATTTCGTTTCAATCAGTGTGCTTCCCGTGGTAG 905
QY 2759 ATAACTACGATACGGAGGGCTTACCATCTGGCCCCAGTCTGCAATGATACCGCGAGAC 2818
Db 904 ATAACTACGATACGGAGGGCTTACCATCTGGCCCCAGTCTGCAATGATACCGCGAGAC 845
QY 2819 CCACGCTCACGGCTCCAGATTATCAGCAATAAACACGACGCGGAAAGGCGGAGCGC 2878
Db 844 CCACGCTCACGGCTCCAGATTATCAGCAATAAACACGACGCGGAAAGGCGGAGCGC 785
QY 2879 AGAAGTGGTCTCTGCAACTTTATCCGCTCCATCCAGTCTATTAAATTTGTCGGGGAAGCT 2938
Db 784 AGAAGTGGTCTCTGCAACTTTATCCGCTCCATCCAGTCTATTAAATTTGTCGGGGAAGCT 725
QY 2939 AGAAGTGGTCTCTGCAACTTTATCCGCTCCATCCAGTCTATTAAATTTGTCGGGGAAGCT 2998
Db 724 AGAAGTGGTCTCTGCAACTTTATCCGCTCCATCCAGTCTATTAAATTTGTCGGGGAAGCT 665
QY 2999 GTGGTGTACGCTCGCTGGTGGTATGGCTTCAATTCAGCTCCGGTCCCAACGATCAAGG 3058
Db 664 GTGGTGTACGCTCGCTGGTGGTATGGCTTCAATTCAGCTCCGGTCCCAACGATCAAGG 605
QY 3059 CGAGTTACATGATCCCGATTTGTGCAAAAAAGCGTTAGCTCTTCGGTCTCCCGATC 3118
Db 604 CGAGTTACATGATCCCGATTTGTGCAAAAAAGCGTTAGCTCTTCGGTCTCCCGATC 545
QY 3119 GTTGTGAGAAAGTAAAGTGGCGGAGTTATCACTCATGTTATGCGAGCAGCTGCAATAT 3178
Db 544 GTTGTGAGAAAGTAAAGTGGCGGAGTTATCACTCATGTTATGCGAGCAGCTGCAATAT 485
QY 3179 TCTCTTACTGTCTATGCCATCCGTAAGATGCTTTTCTGCTGACTGGTCTGACTCAACCAAG 3238
Db 484 TCTCTTACTGTCTATGCCATCCGTAAGATGCTTTTCTGCTGACTGGTCTGACTCAACCAAG 425
QY 3239 TCATTCTGAGATAGTGTATGCGGCAACCGAGTTGTCTTTGCGCGCGCTCAATACCGGAT 3298
Db 424 TCATTCTGAGATAGTGTATGCGGCAACCGAGTTGTCTTTGCGCGCGCTCAATACCGGAT 365
QY 3299 AATACCGCGCCATAGCAGAACTTTAAAGTCTCATATTTGAAAGAAAGCTTTTCGGGG 3358
Db 364 AATACCGCGCCATAGCAGAACTTTAAAGTCTCATATTTGAAAGAAAGCTTTTCGGGG 305
QY 3359 CGAAACTCTCAGAGATCTTACCGCTGTGAGATCCAGTTCGATTAACCCACTCTGCA 3418
Db 304 CGAAACTCTCAGAGATCTTACCGCTGTGAGATCCAGTTCGATTAACCCACTCTGCA 245
QY 3419 CCCAACTGATCTTTCAGCATCTTTTACTTTTCCAGCGGTTTCTGGGTGAGCAAAACAGCA 3478
Db 244 CCCAACTGATCTTTCAGCATCTTTTACTTTTCCAGCGGTTTCTGGGTGAGCAAAACAGCA 185
QY 3479 AGSCAAATGCGCCAAAGAGGAATAAGGGCGACACGGAATGTTGAAATCTCATACTC 9538

Db 184 AGGCAAAATGCCGCAAAAAAGGAATAAGGGCGACACGGAATGTTGAATCTACTACTC 125
QY 3539 TTCCTTTTTCAATATTTATTGAAGCAATTTATCAGGGTTATTGTCTCATGAGCGGATACATA 3598
Db 124 TTCCTTTTTCAATATTTATTGAAGCAATTTATCAGGGTTATTGTCTCATGAGCGGATACATA 65
QY 3599 TTTGAATGTTATTAGAAAAATAAACAAATAGGGTTCCGCGCACATTTCCCGCAAAAGTG 3658
Db 64 TTTGAATGTTATTAGAAAAATAAACAAATAGGGTTCCGCGCACATTTCCCGCAAAAGTG 5
QY 3659 CCAC 3662
Db 4 CCAC 1

RESULT 15
ADS17191/c
ID ADS17191 standard; DNA; 3637 BP.
XX
AC ADS17191;
XX
DT 02-DEC-2004 (first entry)
XX
DE PBSWTGFP plasmid DNA encoding A. victoria GFP Cycle 3 ORF.
XX
KW Polymerase; GFP; jellyfish; green fluorescent protein; chimeric; gene;
XX
OS Aequorea victoria.
XX
OS Chimeric.
XX
OS Unidentified.
XX
PN US2004180352-A1.
XX
PD 16-SEP-2004.
XX
PF 08-AUG-2003; 2003US-00637758.
XX
PR 08-AUG-2002; 2002US-0402342P.
XX
XX (LARG-) LARGE SCALE BIOLOGY CORP.
PI Padgett HS, Lindbo JA, Fitzmaurice WP, Vaewhongs AA;
XX WPI; 2004-667658/65.
XX
XX In vitro method of making sequence variants from heteroduplex
PT polynucleotide, involves combining polynucleotide with agent having
PT polymerase and strand cleavage activities for sufficient time for
PT percentage of complementarity to increase.
XX
PS Example 18; SEQ ID NO 31; 81pp; English.
XX
CC The invention relates to an in vitro method of making sequence variants
CC from heteroduplex polynucleotide which involves combining polynucleotide
CC with an agent having polymerase and strand cleavage activities for
CC sufficient time for the percentage of complementarity to be increased
CC within the heteroduplex. The method is useful for performing in vitro
CC method of making sequence variants from one or more heteroduplex
CC polynucleotide. It is useful in increasing diversity in a population of
CC sequences and obtaining a polynucleotide encoding a desired functional
CC property. The present sequence is a plasmid DNA encoding Aequorea
CC victoria green fluorescent protein (GFP) Cycle' 3 ORF DNA. This sequence
CC is used to illustrate the method of the invention.
XX
SQ Sequence 3637 BP; 959 A; 879 C; 868 G; 931 T; 0 U; 0 Other;

Query Match 72.5%; Score 2656.4; DB 13; Length 3637;
Best Local Similarity 84.0%; Pred. No. 0;
Matches 3079; Conservative 0; Mismatches 556; Indels 29; Gaps 6;
QY 1 CTAATGTTAAGCGTTAATATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTC 60

QY 2219 CGTCTTGAGTCCAAACCGGTAAAGACAGCACTTATCGCACTGCGCACTGCGCACTGCTAAC 2278
DB 1444 CGTCTTGAGTCCAAACCGGTAAAGACAGCACTTATCGCACTGCGCACTGCGCACTGCTAAC 1385
QY 2279 AGGATTAGCAGACGCGAGGTATGTAGGCGGTCTACAGAGTTCCTTGAAGTGGTGGCCTAAC 2338
DB 1384 AGGATTAGCAGACGCGAGGTATGTAGGCGGTCTACAGAGTTCCTTGAAGTGGTGGCCTAAC 1325
QY 2339 TAGCGGTACACTAGAGAGCAGTATTGTTGTPATCTGCGCTCTGCTGAAGCCAGTTACTCTTC 2398
DB 1324 TAGCGGTACACTAGAGAGCAGTATTGTTGTPATCTGCGCTCTGCTGAAGCCAGTTACTCTTC 1265
QY 2399 GGAAGAGAGTGGTAGTCTTTCATCGGCAACCAACCAACCGCTGGTAGCGGTGGTTTT 2458
DB 1264 GGAAGAGAGTGGTAGTCTTTCATCGGCAACCAACCAACCGCTGGTAGCGGTGGTTTT 1205
QY 2459 TTTGTTTGAAGCAGCAGATTACCGCGCAGAAAAAGGATCTCAAGAGATCTCTTTGATC 2518
DB 1204 TTTGTTTGAAGCAGCAGATTACCGCGCAGAAAAAGGATCTCAAGAGATCTCTTTGATC 1145
QY 2519 TTTTCTACGGGGTCTGACGCTCAGTGGAAAGCAAACTCAGTTAAGGGAATTTTGGTCATG 2578
DB 1144 TTTTCTACGGGGTCTGACGCTCAGTGGAAAGCAAACTCAGTTAAGGGAATTTTGGTCATG 1085
QY 2579 AGATTATCAAAAAGGATCTTCACTAGATCTCTTTAAATTTAAATTTAAATCA 2638
DB 1084 AGATTATCAAAAAGGATCTTCACTAGATCTCTTTAAATTTAAATTTAAATCA 1025
QY 2639 ATCTAAGTATATATCAGTAAACCTGCTGACAGTTTACCAATGCTTTAATCAGTAGGCA 2698
DB 1024 ATCTAAGTATATATCAGTAAACCTGCTGACAGTTTACCAATGCTTTAATCAGTAGGCA 965
QY 2699 CCTATCTCAGCGATCTGTCTATTTCCTTCAATCACTAGTTGCTGACTCCCGCTCGTGTAG 2758
DB 964 CCTATCTCAGCGATCTGTCTATTTCCTTCAATCACTAGTTGCTGACTCCCGCTCGTGTAG 905
QY 2759 ATAACTACAGTACGGAGGGCTTACCATCTGGCCCCAGTGTCCATGATACCGCGAGAC 2818
DB 904 ATAACTACAGTACGGAGGGCTTACCATCTGGCCCCAGTGTCCATGATACCGCGAGAC 845
QY 2819 CCACGCTCACGGCTCCAGATTTCATCAGCAATAAACAGCAGCCAGCGAAGGGCCGAGCGC 2878
DB 844 CCACGCTCACGGCTCCAGATTTCATCAGCAATAAACAGCAGCCAGCGAAGGGCCGAGCGC 785
QY 2879 AGAAGTGGTCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTTGCCGGAAGCT 2938
DB 784 AGAAGTGGTCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTTGCCGGAAGCT 725
QY 2939 AGAGTAAGTAGTTGCGCAGTTAATAGTTTGGCGAACGTTGTTGCCATTTGCTACAGGCATC 2998
DB 724 AGAGTAAGTAGTTGCGCAGTTAATAGTTTGGCGAACGTTGTTGCCATTTGCTACAGGCATC 665
QY 2999 GTGGTCTCAGCTCGTCTGTTGGTATGGCTTCAATTCAGTCCGGTCCCAACGATCAAG 3058
DB 664 GTGGTCTCAGCTCGTCTGTTGGTATGGCTTCAATTCAGTCCGGTCCCAACGATCAAG 605
QY 3059 CGAGTTACATGATCCCGCTTGTGTGCAAAAAGCGTTAGTCTCTTCGTTCTCCGATC 3118
DB 604 CGAGTTACATGATCCCGCTTGTGTGCAAAAAGCGTTAGTCTCTTCGTTCTCCGATC 545
QY 3119 GTTGTGAGAAAGTATGGCCGAGTGTATATCACTCATGTTATGGCAGCATGTCATAAT 3178
DB 544 GTTGTGAGAAAGTATGGCCGAGTGTATATCACTCATGTTATGGCAGCATGTCATAAT 485
QY 3179 TCTCTTACATGTCATGCGATCCGTAAGATGCTTTCTGTGACTGGTACTCAACCAAG 3238
DB 484 TCTCTTACATGTCATGCGATCCGTAAGATGCTTTCTGTGACTGGTACTCAACCAAG 425
QY 3239 TCATTCTGAGAAATAGTGTATGGCGGACCGAGTGTCTTGTGCGGCGTCAATACGGGAT 3298
DB 424 TCATTCTGAGAAATAGTGTATGGCGGACCGAGTGTCTTGTGCGGCGTCAATACGGGAT 365
QY 3299 AATACCGCGCCACATAGCAAACTTTAAAGTGTCTCATCATTTGGAAAAAGCTTCTTCGGGG 3358

DB 364 AATACCGCGCCACATAGCAAACTTTAAAGTGTCTCATATTGGAAAAAGCTTCTTCGGGG 305
QY 3359 CGAAAACTCTCAAGGATCTTACCGCTGTGTGAGATCCAGTTCAGATGTAAACCCACTCGTGCA 3418
DB 304 CGAAAACTCTCAAGGATCTTACCGCTGTGTGAGATCCAGTTCAGATGTAAACCCACTCGTGCA 245
QY 3419 CCCAACTGATCTTCAGCATCTTTTACTTTTCAACGAGTTCCTGGGTGAGCAAAAAACAGGA 3478
DB 244 CCCAACTGATCTTCAGCATCTTTTACTTTTCAACGAGTTCCTGGGTGAGCAAAAAACAGGA 185
QY 3479 AGGCAAAATGCCGCAAAAAGGGAATAAGGGCGACACGGAATGTTGAATACTCATACTC 3538
DB 184 AGGCAAAATGCCGCAAAAAGGGAATAAGGGCGACACGGAATGTTGAATACTCATACTC 125
QY 3539 TTTCTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATA 3598
DB 124 TTTCTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATA 65
QY 3599 TTTGAATGTTATTAGAAAAATAAAATAAGGGTTCGCGCACATTTCCCCGAAAAAGTG 3658
DB 64 TTTGAATGTTATTAGAAAAATAAAATAAGGGTTCGCGCACATTTCCCCGAAAAAGTG 5
QY 3659 CCAC 3662
DB 4 CCAC 1

Search completed: June 5, 2005, 14:16:59
Job time : 1798.75 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 5, 2005, 13:39:19 ; Search time 610.194 Seconds
(without alignments)
9819.908 Million cell updates/sec

Title: US-10-826-523-41
Perfect score: 3662
Sequence: 1 ctaaatgaagcgttaata.....atttccccgaagaagtgccac 3662

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfileseq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2639	72.1	3681	1	US-07-924-028A-6
2	2623.6	71.6	5314	1	US-07-924-028A-4
3	2620.6	71.6	3699	3	US-08-846-538-6
4	2620.6	71.6	3699	3	US-09-503-222-6
5	2469.4	67.4	5277	4	US-09-890-848-16
6	2448.8	66.9	3956	4	US-09-402-266B-21
7	2305.2	62.9	4102	4	US-09-402-266B-20
8	2302	62.9	4088	4	US-08-402-266B-18
9	2279.2	62.2	4119	4	US-09-993-170-1
10	2172.8	59.3	7102	3	US-09-138-024-20
11	2172.8	59.3	7102	3	US-09-404-066-20
12	2172.8	59.3	7102	3	US-09-573-322-20
13	2172.8	59.3	7333	3	US-09-138-024-21
14	2172.8	59.3	7333	3	US-09-404-066-21
15	2172.8	59.3	7333	3	US-09-573-322-21
16	2170.6	59.3	6206	2	US-08-474-169-3
17	2170.6	59.3	6206	2	US-08-474-169-2
18	2170.2	59.3	2973	4	US-09-402-266B-17
19	2170.2	59.3	3789	3	US-09-075-019-8
20	2170.2	59.3	4583	4	US-09-402-266B-19
21	2170	59.3	2961	3	US-08-446-935-6
22	2170	59.3	4145	3	US-08-651-472-62
23	2170	59.3	4145	3	US-08-358-928-62
24	2170	59.3	4277	3	US-08-651-472-63
25	2170	59.3	4277	3	US-08-358-928-63
26	2170	59.3	5356	3	US-08-446-935-1
27	2170	59.3	5352	3	US-08-651-472-72

C 28	2170	59.3	5532	3	US-08-358-928-72	Sequence 72, Appl
C 29	2170	59.3	5534	1	US-08-452-267-3	Sequence 3, Appl
C 30	2170	59.3	5534	3	US-09-123-644-3	Sequence 3, Appl
C 31	2170	59.3	6811	3	US-08-651-472-67	Sequence 67, Appl
C 32	2170	59.3	6811	3	US-08-358-928-67	Sequence 67, Appl
C 33	2170	59.3	6926	3	US-08-651-472-69	Sequence 69, Appl
C 34	2170	59.3	6926	3	US-08-358-928-69	Sequence 69, Appl
C 35	2170	59.3	7090	4	US-09-714-550-18	Sequence 18, Appl
C 36	2170	59.3	7214	4	US-09-438-268-1	Sequence 1, Appl
C 37	2170	59.3	7432	4	US-09-367-891A-6	Sequence 6, Appl
C 38	2170	59.3	8560	3	US-09-936-572-11	Sequence 11, Appl
C 39	2170	59.3	11233	3	US-08-980-832-27	Sequence 27, Appl
C 40	2170	59.3	11233	4	US-09-920-923B-27	Sequence 27, Appl
C 41	2169	59.2	3737	2	US-08-784-208-1	Sequence 1, Appl
C 42	2169	59.2	3796	4	US-09-470-661A-32	Sequence 32, Appl
C 43	2169	59.2	4399	4	US-09-911-781-27	Sequence 27, Appl
C 44	2169	59.2	4399	4	US-10-400-902-27	Sequence 27, Appl
C 45	2169	59.2	6295	2	US-08-659-206A-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1

US-07-924-028A-6
; Sequence 6, Application US/07924028A
; Patent No. 5470573

; GENERAL INFORMATION:
; APPLICANT: Lubitz Werner, Szostak, Michael P.

; TITLE OF INVENTION: CARRIER-BOUND RECOMBINANT PROTEINS, PROCESS
; FOR THE PRODUCTION AND USE AS IMMUNOGENS AND VACCINES

; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue

; CITY: New York City
; STATE: New York

; COUNTRY: USA
; ZIP: 10022

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage

; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS

; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/924,028A
; FILING DATE: 30-SEP-1992

; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/EP91/00308
; FILING DATE: 02-FEB-1991

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 40 05 874

; FILING DATE: 24-FEB-1990
; ATTORNEY/AGENT INFORMATION:

; NAME: Hanson, No. 5470573man D.
; REGISTRATION NUMBER: 30,946

; REFERENCE/DOCKET NUMBER: HUBR 1027
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884

; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:

; LENGTH: 3681 base pairs
; TYPE: nucleic acid

; STRANDEDNESS: single
; TOPOLOGY: linear

US-07-924-028A-6

Query Match 72.1%; Score 2639; DB 1; Length 3681;
Best Local Similarity 83.5%; Pred. No. 0;
Matches 3075; Conservative 0; Mismatches 580; Indels 28; Gaps 6;

QY 3 AAATTGTAAAGCGTTAATATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTCAT 62
Db |||||
QY 1 AAATTGTAACGTTAATATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTCAT 60
Db |||||
QY 63 TTTTAAACCAATAGCGCGAAATTCGGCAAAATCCCTTATAAATCAAAAGAAATAGACCGAGA 122
Db |||||
QY 61 TTTTAAACCAATAGCGCGAAATTCGGCAAAATCCCTTATAAATCAAAAGAAATAGACCGAGA 120
Db |||||
QY 123 TAGGGTTGAGTGTGTTTCAGTTTGGAAACAGAGTCCACTATTAAAGAACGTGGACTCCA 182
Db |||||
QY 121 TAGGGTTGAGTGTGTTTCAGTTTGGAAACAGAGTCCACTATTAAAGAACGTGGACTCCA 180
Db |||||
QY 183 AGCTCAAAAGGGGAAACACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCT 242
Db |||||
QY 181 ACCTCAAAAGGGGAAACACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCT 240
Db |||||
QY 243 AATCAAGTTTTTTGGGGTTCAGGTGCGTTAAAGCACTAAATCGBAACCCCTAAAGGGAGCC 302
Db |||||
QY 241 AATCAAGTTTTTTGGGGTTCAGGTGCGTTAAAGCACTAAATCGBAACCCCTAAAGGGAGCC 300
Db |||||
QY 303 CCCGATTTAGAGCTTGACGGGAAAGCCGCGAAAGCTGCGAGAAAGGAGGAAAGAAAG 362
Db |||||
QY 301 CCCGATTTAGAGCTTGACGGGAAAGCCGCGAAAGCTGCGAGAAAGGAGGAAAGAAAG 360
Db |||||
QY 363 CGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGTCACTGCGCGTAAACACCA 422
Db |||||
QY 361 CGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGTCACTGCGCGTAAACACCA 420
Db |||||
QY 423 CACCGCGCGCTTAATGCGCGCTACAGGGGCGTCCCATTCGCGCATTCAGGCTGCGCA 482
Db |||||
QY 421 CACCGCGCGCTTAATGCGCGCTACAGGGGCGTCCCATTCGCGCATTCAGGCTACGCA 480
Db |||||
QY 483 ACTGTTGGGAAGGGGATCGGTGCGGGCTCTTCGCTATTACGCCAGCTGGCGAAAGGGG 542
Db |||||
QY 481 ACTGTTGGGAAGGGGATCGGTGCGGGCTCTTCGCTATTACGCCAGCTGGCGAAAGGGG 540
Db |||||
QY 543 GATGCTGCAAGGGGATTAAGTTGGGTAAAGCCAGGGTTTTCCAGTCAAGCATTTGTA 602
Db |||||
QY 541 GATGCTGCAAGGGGATTAAGTTGGGTAAAGCCAGGGTTTTCCAGTCAAGCATTTGTA 600
Db |||||
QY 603 AAACAGCGCCAGTGAAGCGCGCTCGTTCAITTCAGTTTTTGAACCGGTGAGGAGCGGC 662
Db |||||
QY 601 AAACAGCGCCAGTGAAGTTGTAATACGACTACTATAGGGCGAATTTGGAGCTTCCACCGCG 660
Db |||||
QY 663 AGACTCGCGTGCATAATGTTTTACAGCTGCATGGAGCAGATGAAGTGTGCGACGC 722
Db |||||
QY 661 GTGGCGCGCTCTAGTATGTTGCTACTCTAGTACAATCTGCTCTGATGCCCATAGTTA 720
Db |||||
QY 723 TGCAGAACACGCGCTAGATTAAACCTAGAAAGATAATCATATTGTGACGTACGTAAAG 782
Db |||||
QY 721 AGCCAGTATATACCTCCGCTATCGCTACGTGACTGGGTCTGCTGCCCGGCACGCC 780
Db |||||
QY 783 ATAAATCATGCGTAAATTTGACCATGGGATCTGTAATAGCACTCACTATAGGGCGAATTG 842
Db |||||
QY 781 GCCAACACCGCTGACGCGCCCTGACGGGCTGTCTGCTCCCGGCATCCGCTTACAGACA 840
Db |||||
QY 843 GGTACCGGGCCCCCTCGAGGTTCAGCGTATCGA- - - - -TAAGCTTGATATCAATTC 896
Db |||||
QY 841 AGCTGTGACCGTCTCCGGAGCTGCATGTGTGAGCAATTCGTCCTTAAGTAAGCAATTGCTG 900
Db |||||
QY 897 CTGACGCGGGGGATTCACCTAGTTCTAGAGCGGCGCCACCGCGTGGAGCTCCAGCTT 956
Db |||||
QY 901 CGCGAGGCAAGGTTCGATGCTTTGTGAGCAATTCGTCCTTAAGTAAGCAATTGCTG 960
Db |||||
QY 957 TTGTTCCCTTTAGTAGGGTTAATTTAGATCCCAT- - - - -CGTCAATTTTACGAGAC 1009
Db |||||
QY 961 TAAAGTCTGCTGCTGCGGATCACCGCTTCCAGTAGCGACAGAAAGCAATTGATTTGTA 1020
Db |||||
QY 1010 TATCTTTCTAGGGTTAACTACCTGCATCAGGATCATATCTGTCGGGTCTTTTTCGGCT 1069
Db |||||
QY 1021 TTTTCAGAGAAAGATCGCGAGGAAGATCAATACATTAAGAGTTGAATCTTTTGTGCT 1080
Db |||||
QY 1070 CAGTCATCGCCCAAGCTGCGCTATCTGCGCATCGGGGAGGAAGCCCGCTTTTC 1129
Db |||||

Db |||||
QY 1081 TCGACATGGGTAAATCTCATGTGTTGAATGGCCCTAGAGGATCCGGCCAAGCTTGCATGCC 1140
Db |||||
QY 1130 CGCGAGGTTGAAGCGGCATGGAAAGAGTTTCCCGAGGATGACTGCTGCTCATTCACCT .1189
Db |||||
QY 1141 TCGAGGTGCATCTTAGAGGATCCCGGACGCTCGACGCCATTAATTAATGTTTCCGTAAT 1200
Db |||||
QY 1190 TGAGCGAAAACGACACGCTTTACCATGATGATTCGGGAAGGTGTGGCATGCACGCTTTAA 1249
Db |||||
QY 1201 TCAGGCGCTTCCATCATGAGACAGCGCGCTTT- - -GAAATGTTGACGGGATGAACATAATAG 1258
Db |||||
QY 1250 CGGTCAACTGTTTCGTTTCAGGCGCACTGGGATACCAAGTTCGTGCGGCTTTTCCGACACA 1309
Db |||||
QY 1259 CAATGACGCGCAGCAATAAATCAACAGG- - -AGCAGGAAAGCGAGGTATCCACAAAGT 1315
Db |||||
QY 1310 GTTTCGGATGTCAGCGCGCATCAGCAACCGGAACATACCGGCGACAG- - - - - 1363
Db |||||
QY 1316 CCAGCGTACCAATAAACGCAAGCCTCAACGACGCGACGACGAGGCGGTCACTAGCA 1375
Db |||||
QY 1364 CCGGAATCGCGTGGCGGTGTGAGATTAATGACAGCGGTGGCGCTGGGATATTCGT 1423
Db |||||
QY 1376 TCCAAACTTGTACTCTGTCAGAAATCGAATCATCTTCGTTAAATCCAAAACGGCAG 1435
Db |||||
QY 1424 CAGCGAGAACGGGTAT- - - - -CTGGCTGGATGCCCGAAGAAATGGAATGATACCCCGTG 1479
Db |||||
QY 1436 AAGCCTGAATGAGAATTCGACCTCGAGGGGGGCGCGGTACCCAGCTTTTGTCCCTTTA 1495
Db |||||
QY 1480 AGTTACCCCGCGCGCTTGGCGTAAATCATGTTGTCATGCTGTTTCTGTCGAAATG 1539
Db |||||
QY 1496 GTGAGGGTTAATTCGAGCTTGGCGTAAATCATGTTGTCATGCTGTTTCTGTGAAATG 1555
Db |||||
QY 1540 TTATCCGCTCAAAATTCACACAAATACGAGCCGGAAGCAATAAGTGTAAAGCTGGGG 1599
Db |||||
QY 1556 TTATCCGCTCAAAATTCACACAAATACGAGCCGGAAGCAATAAGTGTAAAGCTGGGG 1615
Db |||||
QY 1600 TGCCTAATAGTGAAGTAACTCAATTAATTTGCTGTTGCTCCTGCGCTTTTCCAGTC 1659
Db |||||
QY 1616 TGCCTAATAGTGAAGTAACTCAATTAATTTGCTGTTGCTCCTGCGCTTTTCCAGTC 1675
Db |||||
QY 1660 GGGAAACCTGTCGTCAGCTGCATTAAATGAAATCGGCCAAACGCGCGGGAGAGCGGTTT 1719
Db |||||
QY 1676 GGGAAACCTGTCGTCAGCTGCATTAAATGAAATCGGCCAAACGCGCGGGAGAGCGGTTT 1735
Db |||||
QY 1720 GCGTATTCGGCGCTCTTCGCTCTCCTCACTCACTCACTCGCTCGCTCGCTCGCTCGCT 1779
Db |||||
QY 1736 GCGTATTCGGCGCTCTTCGCTCTCCTCACTCACTCACTCGCTCGCTCGCTCGCTCGCT 1795
Db |||||
QY 1780 GCGCGAGCGGTATCAGCTCACTCAAGCGGTAATACGGTTATCCACAGAAATCAGGGA 1839
Db |||||
QY 1796 GCGCGAGCGGTATCAGCTCACTCAAGCGGTAATACGGTTATCCACAGAAATCAGGGA 1855
Db |||||
QY 1840 TAAACGAGAAAGAACATGTGAGCAAAAGGCGCAAAAGGCGCAGAAACCGTAAAGAGC 1899
Db |||||
QY 1856 TAAACGAGAAAGAACATGTGAGCAAAAGGCGCAGAAAGGCGCAGAAACCGTAAAGAGC 1915
Db |||||
QY 1900 GCGCTGCTGCGCTTTTCCATAGGCTCGCGCCCTCGACGAGCATCAAAAATTCGAGC 1959
Db |||||
QY 1916 GCGTGTGCTGCGCTTTTCCATAGGCTCGCGCCCTCGACGAGCATCAAAAATTCGAGC 1975
Db |||||
QY 1960 CTCAGTCAAGGTGCGCAACCCGACAGGACTATAAGATACACAGGCTTTCCCGCTGG 2019
Db |||||
QY 1976 CTCAGTCAAGGTGCGCAACCCGACAGGACTATAAGATACACAGGCTTTCCCGCTGG 2035
Db |||||
QY 2020 AAGCTCCCTGCTGCGCTCTCTGTTCCGACCTCGCTTACCGGATACCTGTCGCGCTT 2079
Db |||||
QY 2036 AAGCTCCCTGCTGCGCTCTCTGTTCCGACCTCGCTTACCGGATACCTGTCGCGCTT 2095
Db |||||
QY 2080 TCTCCCTTCGCGAAGCTGCGGCTTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGCT 2139
Db |||||
QY 2096 TCTCCCTTCGCGAAGCTGCGGCTTTTCTCAATGCTCAGCTGTAGGTATCTCAGTTCGCT 2155
Db |||||
QY 2140 GTAGTCTGTTCCGCTCAAGCTGGGCTGTGTGCAAGAACCCCGCTTTCAGCCGACCGCTG 2199
Db |||||

Db 2156 GTAGTTCGTTCCGCTCCAGCTGGCTGTGTGACGAACCCCGTTTCAGCCGACCGCTG 2215
Qy 2200 CGCTTTATCCGGTAACTATCTGTTGAGTCCAAACCCGGTAAAGACACGACTTATCGCACT 2259
Db 2216 CGCTTTATCCGGTAACTATCTGTTGAGTCCAAACCCGGTAAAGACACGACTTATCGCACT 2275
Qy 2260 GGCAGGACCACTGGTAAACAGGATTAACGACGAGGAGTGTAGGCGGTCTACAGAGTT 2319
Db 2276 GGCAGGACCACTGGTAAACAGGATTAACGACGAGGAGTGTAGGCGGTCTACAGAGTT 2335
Qy 2320 CTTGAAGTGTGGCTAACTACGCTTACCTAGAGGACAGTATTTGGTATCTCGGCTCT 2379
Db 2336 CTTGAAGTGTGGCTAACTACGCTTACCTAGAGGACAGTATTTGGTATCTCGGCTCT 2395
Qy 2380 GCTGAAGCAGTTACCTTCCGAAAAAGAGTGTGTAGCTCTTGTGTCGGGCAAAACAAACCCAC 2439
Db 2396 GCTGAAGCAGTTACCTTCCGAAAAAGAGTGTGTAGCTCTTGTGTCGGGCAAAACAAACCCAC 2455
Qy 2440 CGCTGGTAGCGGTGGTTTTTCTGTTTGCAGAGCAGCAGATTACGCGCAGAAAAAAGGATC 2499
Db 2456 CGCTGGTAGCGGTGGTTTTTCTGTTTGCAGAGCAGCAGATTACGCGCAGAAAAAAGGATC 2515
Qy 2500 TCAAGAGAGTCTTTGATCTTTCTACGGGTCTGACGCTCAGTGGAAACGAAACTCACG 2559
Db 2516 TCAAGAGAGTCTTTGATCTTTCTACGGGTCTGACGCTCAGTGGAAACGAAACTCACG 2575
Qy 2560 TTAAGGGATTTGGTCATGAGATTATCAAAAGGATCTTCACCTAGATCTCTTTAAATTA 2619
Db 2576 TTAAGGGATTTGGTCATGAGATTATCAAAAGGATCTTCACCTAGATCTCTTTAAATTA 2635
Qy 2620 AAAATGAAGTTTTAAATCAATCTAAAGTATATATAGTAAACTTGGTCTGACAGTTACCA 2679
Db 2636 AAAATGAAGTTTTAAATCAATCTAAAGTATATATAGTAAACTTGGTCTGACAGTTACCA 2695
Qy 2680 ATGCTTAATCAGTAGGACCTATCTCAGCGATCTGTCTATTTCTGTTTCATCCATAGTTGC 2739
Db 2696 ATGCTTAATCAGTAGGACCTATCTCAGCGATCTGTCTATTTCTGTTTCATCCATAGTTGC 2755
Qy 2740 CTGACTCCCCGTCGTGTAGATACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGC 2799
Db 2756 CTGACTGCCGTCGTGTAGATACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGC 2815
Qy 2800 TGAATGATACCGGAGACCCAGCTCACCAGCTCCAGATTATCAGCAATAAACCGACC 2859
Db 2816 TGAATGATACCGGAGACCCAGCTCACCAGCTCCAGATTATCAGCAATAAACCGACC 2875
Qy 2860 AGCCGGAAGGCGGAGCGCAGAGAGTGTCTCGAACTTTATCCGCTCCATCCAGTCTAT 2919
Db 2876 AGCCGGAAGGCGGAGCGCAGAGAGTGTCTCGAACTTTATCCGCTCCATCCAGTCTAT 2935
Qy 2920 TAATTGTTGCCGGAAGCTAGAGTAAGTAGTTGCCAGTTAATAGTTTGCACAACTGTG 2979
Db 2936 TAATTGTTGCCGGAAGCTAGAGTAAGTAGTTGCCAGTTAATAGTTTGCACAACTGTG 2995
Qy 2980 TGCATTTGCTACAGGCAATCGTGTGTCACTGCTGCTGTTGGTATGGCTTCATTCAGCTC 3039
Db 2996 TGCATTTGCTACAGGCAATCGTGTGTCACTGCTGCTGTTGGTATGGCTTCATTCAGCTC 3055
Qy 3040 CGGTTCCACAGATCAAGGCGAGTTACATCATCCCCCATGTTGTCGAAAAAGCGGTAG 3099
Db 3056 CGGTTCCACAGATCAAGGCGAGTTACATCATCCCCCATGTTGTCGAAAAAGCGGTAG 3115
Qy 3100 CTCTTTCCGTCCTCCGATCGTTGTCAAGAGTGTGGCGCAGTGTATCACTCATGTT 3159
Db 3116 CTCTTTCCGTCCTCCGATCGTTGTCAAGAGTGTGGCGCAGTGTATCACTCATGTT 3175
Qy 3160 TATGGCAGCATCGCATAAATCTCTTACTGTCTATGCCCATCCGTAAGATGCTTTCTGTGAC 3219
Db 3176 TATGGCAGCATCGCATAAATCTCTTACTGTCTATGCCCATCCGTAAGATGCTTTCTGTGAC 3235
Qy 3220 TGGTGATCTCAACCAAGTCATTCTGAGATAGTGTATCGGGGACCGAGTTGCTCTTG 3279
Db 3236 TGGTGATCTCAACCAAGTCATTCTGAGATAGTGTATCGGGGACCGAGTTGCTCTTG 3295

Qy 3280 CCGGCGTCAATACGGGATTAATACCGGCCACATAGCAGAACTTTAAAAGTGTCTCATCAT 3339
Db 3296 CCGGCGTCAATACGGGATTAATACCGGCCACATAGCAGAACTTTAAAAGTGTCTCATCAT 3355
Qy 3340 TGCAGAAACGTTCTTTCGGGGGAAAACTCTCAAGATCTTCAAGATCTTACCGCTGTTGAGATCCAGTTC 3399
Db 3356 TGCAGAAACGTTCTTTCGGGGGAAAACTCTCAAGATCTTACCGCTGTTGAGATCCAGTTC 3415
Qy 3400 GATGTAACCCACTCGTGCCACCCCAACTGATCTTCAGCATCTTTTACTTTTCCACGCGTTTC 3459
Db 3416 GATGTAACCCACTCGTGCCACCCCAACTGATCTTCAGCATCTTTTACTTTTCCACGCGTTTC 3475
Qy 3460 TGGTGAGCAAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGAATAAGGGCGACACGGAA 3519
Db 3476 TGGTGAGCAAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGAATAAGGGCGACACGGAA 3535
Qy 3520 ATGTTGAATACTCATACTCTTCTCTTTTCAATATTTATGAAGCATTTTATCAGGGTTATTG 3579
Db 3536 ATGTTGAATACTCATACTCTTCTCTTTTCAATATTTATGAAGCATTTTATCAGGGTTATTG 3595
Qy 3580 TCTCATGAGCGGATACATATTTGAATGTTATTTAGAAAAATAAACAATAGGGGTTCCGCG 3639
Db 3596 TCTCATGAGCGGATACATATTTGAATGTTATTTAGAAAAATAAACAATAGGGGTTCCGCG 3655
Qy 3640 CACATTTCCCGAAAAAGTGCCAC 3662
Db 3656 CACATTTCCCGAAAAAGTGCCAC 3678

RESULT 2

US-07-924-028A-4
; Sequence 4, Application US/07924028A
; Patent No. 5470573
; GENERAL INFORMATION:
; APPLICANT: Lubitz Werner, Szostak, Michael P.
; TITLE OF INVENTION: CARRIER-BOUND RECOMBINANT PROTEINS, PROCESS
; TITLE OF INVENTION: FOR THE PRODUCTION AND USE AS IMMUNOGENS AND VACCINES
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felte & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/924,028A
; FILING DATE: 30-SEP-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP91/00308
; FILING DATE: 02-FEB-1991
; PRIOR APPLICATION DATA: DE 40 05 874
; FILING DATE: 24-FEB-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5470573man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: HUBR 1027
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5314 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

[illegible]

2097 TGCGGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTTCGGTGTAGGTCTGTTTCGTTCCCA 2156
Db TGCGGCTTTCTCAATGCTCAGCTGTAGGTATCTCAGTTTCGGTGTAGGTCTGTTTCGTTCCCA 3805
2157 AGCTGGGCTGTGTGACGAAACCCCGTTTCAGCCGACCGCTGCGCTTTATCCGGTAACT 2216
Db AGCTGGGCTGTGTGACGAAACCCCGTTTCAGCCGACCGCTGCGCTTTATCCGGTAACT 3865
2217 ATCGTCTTGAGTCCAAACCGGTAAAGACAGACTTATCGCCACTGGCAGAGCCACTGGTA 2276
Db ATCGTCTTGAGTCCAAACCGGTAAAGACAGACTTATCGCCACTGGCAGAGCCACTGGTA 3925
2277 ACAGGATTAGCAGAGGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGTGTCCTTA 2336
Db ACAGGATTAGCAGAGGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGTGTCCTTA 3985
2337 ACTACGGCTACACTAGAGGAGATTTGGTATCTGCGCTCTGCTGGAAGCCAGTTACCT 2396
Db ACTACGGCTACACTAGAGGAGATTTGGTATCTGCGCTCTGCTGGAAGCCAGTTACCT 4045
2397 TCGGAAAAAGAGTGTGTAGCTCTTGATCCGGCAAAACAAACCCGCTGGTAGCGTGGTT 2456
Db TCGGAAAAAGAGTGTGTAGCTCTTGATCCGGCAAAACAAACCCGCTGGTAGCGTGGTT 4105
2457 TTTTGTGTGAAAGAGAGATTACCGCGCAAAAAAAGATCTCAAGAGATCTCTTGA 2516
Db TTTTGTGTGAAAGAGAGATTACCGCGCAAAAAAAGATCTCAAGAGATCTCTTGA 4165
2517 TCTTTTCTACGGGCTGTAGCTCAGTGGAAAGAACTCAGTTAAGGATTTTGGTCA 2576
Db TCTTTTCTACGGGCTGTAGCTCAGTGGAAAGAACTCAGTTAAGGATTTTGGTCA 4225
2577 TGAGATTATCAAAAGGATCTTCACCTAGATCTCTTTAAATTAATAAGATTTTAAAT 2636
Db TGAGATTATCAAAAGGATCTTCACCTAGATCTCTTTAAATTAATAAGATTTTAAAT 4285
2637 CAATCTAAAGTATATATAGTAACTTGTGTGACAGTTAACCAATGCTTAAATCAGTGAGG 2696
Db CAATCTAAAGTATATATAGTAACTTGTGTGACAGTTAACCAATGCTTAAATCAGTGAGG 4345
2697 CACCTATCTCAGGATCTGTATTTTGGTTCATCATAGTTGCTGACTCCCGTCTGT 2756
Db CACCTATCTCAGGATCTGTGTATTTTGGTTCATCATAGTTGCTGACTCCCGTCTGT 4405
2757 AGATACTACGATACGGGAGGCTTACCATCTGGCCCGAGTGTGCAATGATACCGGAG 2816
Db AGATACTACGATACGGGAGGCTTACCATCTGGCCCGAGTGTGCAATGATACCGGAG 4465
2817 ACCACGCTCACGGTCCAGATTTATCAGCAATAAACCCAGCCAGCCGGAAGGCGGAGC 2876
Db ACCACGCTCACGGTCCAGATTTATCAGCAATAAACCCAGCCAGCCGGAAGGCGGAGC 4525
2877 GCAGAGTGGTCTGCAACTTATCCGCTCCATCCAGTCTATTATTTGTCGGGAG 2936
Db GCAGAGTGGTCTGCAACTTATCCGCTCCATCCAGTCTATTATTTGTCGGGAG 4585
2937 CTAGAGTAAAGTGTGCGCAGTTAATGTTTGGCAACGTTGTCATGCTTACAGGCA 2996
Db CTAGAGTAAAGTGTGCGCAGTTAATGTTTGGCAACGTTGTCATGCTTACAGGCA 4645
2997 TCGTGGTGTACGCTCGTGTGTGTGATGAGCTTTCATTTCAGTCCGGTCCCAACGATCAA 3056
Db TCGTGGTGTACGCTCGTGTGTGTGATGAGCTTTCATTTCAGTCCGGTCCCAACGATCAA 4705
3057 GGGAGTTACATGATCCCACTGTTGTGCAAAAAAGGTTAGTCTTCCGTCCTCCGA 3116
Db GGGAGTTACATGATCCCACTGTTGTGCAAAAAAGGTTAGTCTTCCGTCCTCCGA 4765
3117 TCGTGTGCAAGTAAAGTTGGCGCAGTGTATCACTCATGTTATGGCAGCACTGCATA 3176
Db TCGTGTGCAAGTAAAGTTGGCGCAGTGTATCACTCATGTTATGGCAGCACTGCATA 4825

3177 ATTCTTCTCTCATGCCATCGTAAGATGCTTTTCTGACTGCTGAGTACTCAACCA 3236
Db ATTCTTCTTCTCTCATGCCATCGTAAGATGCTTTTCTGACTGCTGAGTACTCAACCA 4885
3237 AGTCATTCTGAGAAAGTGTATGCGCGCACCGAGTTGCTCTTGGCCGGGTCAATACGGG 3296
Db AGTCATTCTGAGAAAGTGTATGCGCGCACCGAGTTGCTCTTGGCCGGGTCAATACGGG 4945
3297 ATAATACCGCGCACATAGCAGAACTTTAAAAAGTGTCTCATTTGGAACCGTTCTTCGG 3356
Db ATAATACCGCGCACATAGCAGAACTTTAAAAAGTGTCTCATTTGGAACCGTTCTTCGG 5005
3357 GCGGAAACCTCTCAAGGATCTTACCGCTCTGAGATCCAGTTCGATGTAACCCACTCGTG 3416
Db GCGGAAACCTCTCAAGGATCTTACCGCTCTGAGATCCAGTTCGATGTAACCCACTCGTG 5065
3417 CACCCAACTGATCTTTCAGCATCTTTTACCTTTTACCCAGCGTTTCTGGGTGAGCAAAACAG 3476
Db CACCCAACTGATCTTTCAGCATCTTTTACCTTTTACCCAGCGTTTCTGGGTGAGCAAAACAG 5125
3477 GAAGCAAAATGCGCAAAAAAGGAAATGAAGGCGCACGGAATGTTGAATCTCATAC 3536
Db GAAGCAAAATGCGCAAAAAAGGAAATGAAGGCGCACGGAATGTTGAATCTCATAC 5185
3537 TCTTCTCTTTTCAATATTTTGAAGCATTTATCAGGTTTATTCAGGTTTATTCATGAGCGGATACA 3596
Db TCTTCTCTTTTCAATATTTTGAAGCATTTATCAGGTTTATTCAGGTTTATTCATGAGCGGATACA 5245
3597 TATTTGAATGTTTATGAAAAATTAACAAATAGGGGTTCCGCGCACATTTCCCGGAAAG 3656
Db TATTTGAATGTTTATGAAAAATTAACAAATAGGGGTTCCGCGCACATTTCCCGGAAAG 5305
3657 TGCCAC 3662
Db TGCCAC 5311

RESULT 3
US-08-646-538-6
; Sequence 6, Application US/08646538
; Patent No. 6027881
; GENERAL INFORMATION:
; APPLICANT: Pavlakis, George N.
; APPLICANT: Gaitanaris, George A.
; APPLICANT: Stauber, Roland H.
; APPLICANT: Vournakis, John N.
; TITLE OF INVENTION: Mutant Aequorea victoria Fluorescent
; TITLE OF INVENTION: Proteins Having Increased Cellular Fluorescence
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/646,538
; FILING DATE: No. 6027881 yet assigned
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Kenneth A.
; REGISTRATION NUMBER: 31,677
; REFERENCE/DOCKET NUMBER: 015280-249000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:	
LENGTH: 3699 base pairs	
TYPE: nucleic acid	
STRANDEDNESS: single	
TOPOLOGY: linear	
MOLECULE TYPE: DNA	
FEATURE:	
NAME/KEY: -	
LOCATION: 1..3699	
OTHER INFORMATION: /note= "pBSGFP"	
US-08-646-538-6	
Query Match 71.6%; Score 2620.6; DB 3; Length 3699;	
Best Local Similarity 83.3%; Pred. No. 0;	
Matches 3078; Conservative 0; Mismatches 579; Indels 40; Gaps 7;	
QY 3 AAATTGTAAGCGTTAAATATTTTGTAAATTCGCGTTAAATTTTGTAAATTCAGCTCAT 62	Db
DB 3 AAATTGTAACGTTAAATATTTTGTAAATTCGCGTTAAATTTTGTAAATTCAGCTCAT 62	QY
QY 63 TTTTAAACCAATAGCGCGAAATCGGCAAAATCCCTTATAAATCAAAAGAAATAGACCGAGA 122	Db
DB 63 TTTTAAACCAATAGCGCGAAATCGGCAAAATCCCTTATAAATCAAAAGAAATAGACCGAGA 122	QY
QY 123 TAGGGTTGAGTGTGTTCAGTTTGGAAACAAGAGTCCACTATTAAAGAAACGTGGACTCCA 182	Db
DB 123 TAGGGTTGAGTGTGTTCAGTTTGGAAACAAGAGTCCACTATTAAAGAAACGTGGACTCCA 182	QY
QY 183 AGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTAGCTGAACCATCACCTT 242	Db
DB 183 AGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTAGCTGAACCATCACCTT 242	QY
QY 243 AATCAAGTTTTTGGGTCGAGTGCCTGTAAGAGCACTAAATCGGACCCCTTAAGGGAGCC 302	Db
DB 243 AATCAAGTTTTTGGGTCGAGTGCCTGTAAGAGCACTAAATCGGACCCCTTAAGGGAGCC 302	QY
QY 303 CCCGATTTAGAGCTTTGACGGGAAACCGCGGAAACGTCGCGAAGAAAGGAAAGAAAG 362	Db
DB 303 CCCGATTTAGAGCTTTGACGGGAAACCGCGGAAACGTCGCGAAGAAAGGAAAGAAAG 362	QY
QY 363 CGAAAGGAGCGGGCGTAGGGCGTGGCAAGTGTAGCGGTCA CGCTGCGGTAAACCA 422	Db
DB 363 CGAAAGGAGCGGGCGTAGGGCGTGGCAAGTGTAGCGGTCA CGCTGCGGTAAACCA 422	QY
QY 423 CACCCGCGCGCTTAATGCGCGCTACAGGGCGCGTC -- CCATTGCGCATTCAGCTGCG 479	Db
DB 423 CACCCGCGCGCTTAATGCGCGCTACAGGGCGCGTCGGCCCATTCGCCATTCAGCTGCG 482	QY
QY 480 GCAACTGTTGGGAAGGCGCATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAG 539	Db
DB 480 GCAACTGTTGGGAAGGCGCATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAG 539	QY
QY 540 GGGGATGTCTGCAAGGCGATTAAGTTGGGTAAACCGCAGGGTTTTCCAGTCACGACGTT 599	Db
DB 540 GGGGATGTCTGCAAGGCGATTAAGTTGGGTAAACCGCAGGGTTTTCCAGTCACGACGTT 599	QY
QY 602 GTAAACCGACCGCCAGTGAATTTGTAAT -- ACACCTACTATAGGGCGAAATGGGTACCGG 660	Db
DB 602 GTAAACCGACCGCCAGTGAATTTGTAAT -- ACACCTACTATAGGGCGAAATGGGTACCGG 660	QY
QY 660 GGCAGACTCGCGGTGCAATGTGTTTATACAGCGTGAATGAGAGATGAAGATGCTCGACA 719	Db
DB 660 GGCAGACTCGCGGTGCAATGTGTTTATACAGCGTGAATGAGAGATGAAGATGCTCGACA 719	QY
QY 720 CGCTGCA -- GAACACGCGAGCTAGATTAACCCCTAGAAAGATAATCATATGTGACGTACG 776	Db
DB 720 CGCTGCA -- GAACACGCGAGCTAGATTAACCCCTAGAAAGATAATCATATGTGACGTACG 776	QY
QY 777 TTAAGAGATAATCATCGTAAATATTGACGCATGGGATCTGTAATAGACTCACTATAGGGC 836	Db
DB 777 TTAAGAGATAATCATCGTAAATATTGACGCATGGGATCTGTAATAGACTCACTATAGGGC 836	QY
QY 839 TTCGTTGGGATCTTTTGAAGGGGAGATGTGTGACACAGGTA - ATGGTTGTCTGGTAA 896	Db
DB 839 TTCGTTGGGATCTTTTGAAGGGGAGATGTGTGACACAGGTA - ATGGTTGTCTGGTAA 896	QY

840	GGACAGGGCCATCGCCAAATTTGGAGTATTTTGTGATAATTTGGTCTGCTAGTTGAACGCTTC	899
897	CTGACAGCCCGGGGATCCACTAGTTT --- CTAGACGGCGCGCCACCGCGGTGGAGTCCA	952
900	CATCTTCAATGTGTCTAATTTTGAAGTTAACTTTGATTCCTTTTGTGTCTG	959
953	GCCTTTGTTCCCTTTAGTGAGGGTTAAITAGATCCCATCGCTCAATTTTACCCAGACTAT	1012
960	CCATGATGTATACATTGTGTGAGTTATAGTTGATTTCCAAATTTGTGTTCCAAAGATTTTC	1019
1013	CTTCTAGGGTTAATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGCGCTCAG	1072
1020	CATCTTCTTTAAATCAATACCTTTTAACTCGATTCTATTAAACAAAGGTATCACCTTCAA	1079
1073	TCATCGCCCAAGCTCGCGCTATCTCGGCATCGGGAGGAAGGACCGCTTCCTTTCCCG	1132
1080	ACTTGACTTCAGCACGCTGTCTTGTAGTTCCCGTCACTTTTGAAAAATATAGTTCCTTCT	1139
1133	CGAGGTTGAAGCGGCATCGAAAGAGTTTTCGCGAGATGACTGCTGCTGCTGATTTGA	1192
1140	GTAATTAACCTTCGCGGCATGGCACTCTTTGAAAAAGTCAATCGCGTTTCATATGATCGGGT	1199
1193	CGAAAAACGACGTTTACCATGATGATTCGGGAAGGTGTGGCCATGCAACGCTTTTAAACGG	1252
1200	ATCTTTGAAAAAGCATTTGAACACCCATAAGAGAAAGTAGTGACAAGTGTGGCCATGGAACAG	1259
1253	TGAACGTGTTCGTTAGGGCCACCTGGGATACCAAGTTTCGCGGCTTTTCCGACACAGTT	1312
1260	GTAGTTTTCAGTAGTGTGCAAAATAAATTTAAGGGTAAAGTTTTCGATGTGTGATCACCTT	1319
1313	CGGATGTGTAGCGCGAAGCGCATCAGCAACCCGAAACAATACCGCGCAGCAGCCGGAACGTG	1372
1320	CACCTCTCCACTGACAGAAATTTGTGCCCATTTAAGCATCACCATCTAATTTCAACAGAA	1379
1373	CGTGCC --- GGTGTGCAGATTAATGACAGCGGTGCGGCGTGGGATAT	1418
1380	TTGGGACAACCTCAGTGAAGAGTTCTTCTCTTTCTAGCCATTTCTTTCGCGATCGAAT	1439
1419	TACGTACAGGAGCGGTATCTCTGCTGGATGCGCGAGAAATGACATGGATACCCGT	1478
1440	TCCTGACCGCGGGGATCCACTAGTTCTAGAGCGGCCCCACCGCGGTGGAGTCCACG	1499
1479	GAGTTACCCCGCGG --- CGCGCTTTGGGCTAATCATATGCTCATAGCTGTTT	1525
1500	TTTTGTTCCCTTAGTGAGGGTTAATTCGAGCTTGGCGTAATCATGCTCATAGCTGTTT	1559
1526	CCTGTGAAATTTGTTATCCGCTCAAAATTCACAAACATACAGAGCCGGAAGCATAAAG	1585
1560	CCTGTGAAATTTGTTATCCGCTCAAAATTCACAAACATACAGAGCCGGAAGCATAAAG	1619
1586	TGTAAAGCTGGGTGCTTAATGATGAGCTAACTCACAATTAATTTGCTGCTGCTCACCTG	1645
1620	TGTAAAGCTGGGTGCTTAATGATGAGCTAACTCACAATTAATTTGCTGCTGCTCACCTG	1679
1646	CCCGCTTTCCAGTCGGGAAACCTGTGTCGACAGCTGCAATTAATGAATCGGCCAAACGCGG	1705
1680	CCCGCTTTCCAGTCGGGAAACCTGTGTCGACAGCTGCAATTAATGAATCGGCCAAACGCGG	1739
1706	GGGAGAGCGGTTTGTGTTATTTGGGCGCTTTCGCTTCTCGCTCATGCTGCTGCGC	1765
1740	GGGAGAGCGGTTTGTGTTATTTGGGCGCTTTCGCTTCTCGCTCATGCTGCTGCGC	1799
1766	TCGCTGCTTTCGCTGCGGAGCGGTATCAGCTCACTCAAAAGCGGTAAATACGGTTATCC	1825
1800	TCGCTGCTTTCGCTGCGGAGCGGTATCAGCTCACTCAAAAGCGGTAAATACGGTTATCC	1859
1826	ACAGAAATCAGGGATTAACCGCAGGAAAGAAACATGTGAGCAAAAGGCGCAGAAAGCCAGG	1885
1860	ACAGAAATCAGGGATTAACCGCAGGAAAGAAACATGTGAGCAAAAGGCGCAGAAAGCCAGG	1919
1886	AACCGTAAAGGCGCGGTTGCTGCGGTTTTTCCATAGGCTCCGCGCCCTCGACGAGAT	1945

Db 1920 AACGTAAGAGGCGCGTGTCTGGCGTTTTTCCATAGGCTCCGCCCTCGAGGACAT 1979
Qy 1946 CACAAAAATCGACGCTCAAGTCAAGGTGCGAAACCCGACAGGACTATTAAGATACCAG 2005
Db 1980 CACAAAAATCGACGCTCAAGTCAAGGTGCGAAACCCGACAGGACTATTAAGATACCAG 2039
Qy 2006 GCGTTTCCCGCTCGAAGCTCCCTCGTGGCTCTCTCTGTTCCGACCCCTGCGCTTACCGGA 2065
Db 2040 GCGTTTCCCGCTCGAAGCTCCCTCGTGGCTCTCTCTGTTCCGACCCCTGCGCTTACCGGA 2099
Qy 2066 TACCTGTCCCGCTTCTCCCTTCGCGAAGCGTGGCGTTTCTCATAGCTCACGCTGTAGG 2125
Db 2100 TACCTGTCCCGCTTCTCCCTTCGCGAAGCGTGGCGTTTCTCATAGCTCACGCTGTAGG 2159
Qy 2126 TATCTCAGTTCCGTTGAGTGTGCTTCCGCTCAAGCTGGCTGTGTGCAACGAAACCCCGCTT 2185
Db 2160 TATCTCAGTTCCGTTGAGTGTGCTTCCGCTCAAGCTGGCTGTGTGCAACGAAACCCCGCTT 2219
Qy 2186 CAGCCCGACCGCTCGCGCTTATCCGTTAACTATCTGTTGAGTCCAAACCGGTAAGACAC 2245
Db 2220 CAGCCCGACCGCTCGCGCTTATCCGTTAACTATCTGTTGAGTCCAAACCGGTAAGACAC 2279
Qy 2246 GACTTATCGCACTGGCAGCAGCACTGGTAACAGGATTAGCAGAGCGAGTATGTAGGC 2305
Db 2280 GACTTATCGCACTGGCAGCAGCACTGGTAACAGGATTAGCAGAGCGAGTATGTAGGC 2339
Qy 2306 GGTGTACAGAGTTCTTGAAGTGGTGGCTTAACCTACCGCTACACTAGAAAGGACAGTATT 2365
Db 2340 GGTGTACAGAGTTCTTGAAGTGGTGGCTTAACCTACCGCTACACTAGAAAGGACAGTATT 2399
Qy 2366 GGTATCTGGCTCTGTGAAGCAGTTACCTTCGGAAGGAGTGTGGTAGCTCTTGATCC 2425
Db 2400 GGTATCTGGCTCTGTGAAGCAGTTACCTTCGGAAGGAGTGTGGTAGCTCTTGATCC 2459
Qy 2426 GGCACAAACACACCGCTGTGTAGCGTGGTTTTTTTGTGTTGCAAGCAGCAGATTAGCGC 2485
Db 2460 GGCACAAACACACCGCTGTGTAGCGTGGTTTTTTTGTGTTGCAAGCAGCAGATTAGCGC 2519
Qy 2486 AGAAAAAAGGATCTCAAGAAGATCTCTTTGATCTTTTCTACGGGCTGTGACGCTCAGTGG 2545
Db 2520 AGAAAAAAGGATCTCAAGAAGATCTCTTTGATCTTTTCTACGGGCTGTGACGCTCAGTGG 2579
Qy 2546 AACGAAACTCAGTTAAGGATTTTGGTCATGAGATTATCAAAAAGGATCTTACCTAG 2605
Db 2580 AACGAAACTCAGTTAAGGATTTTGGTCATGAGATTATCAAAAAGGATCTTACCTAG 2639
Qy 2606 ATCTTTTAAATTAATAAGTGTAAATCAATCTAAAGTATATATGAGTAAACTTGG 2665
Db 2640 ATCTTTTAAATTAATAAGTGTAAATCAATCTAAAGTATATATGAGTAAACTTGG 2699
Qy 2666 TCTGACAGTTACCAATGCTTAAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGT 2725
Db 2700 TCTGACAGTTACCAATGCTTAAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGT 2759
Qy 2726 TCATCATAGTTGCTGATCTCCCGCTGTGTAGATACTACGATACGGAGGCGCTTACCA 2785
Db 2760 TCATCATAGTTGCTGATCTCCCGCTGTGTAGATACTACGATACGGAGGCGCTTACCA 2819
Qy 2786 TCTGGCCCGAGTCTGCAATGATACCGCAGACCCAGCTCACCGCTCCAGATTATCA 2845
Db 2820 TCTGGCCCGAGTCTGCAATGATACCGCAGACCCAGCTCACCGCTCCAGATTATCA 2879
Qy 2846 GCAATAAACCCAGCCGAGCGGCGCAGCGCAGAGTGGTCTTCAACTTTATCCGCC 2905
Db 2880 GCAATAAACCCAGCCGAGCGGCGCAGCGCAGAGTGGTCTTCAACTTTATCCGCC 2939
Qy 2906 TCATCATAGTTAATTTGTTCCCGGAGGCTAGAGTAAAGTGTGCGCAGTTAATAGT 2965
Db 2940 TCATCATAGTTAATTTGTTCCCGGAGGCTAGAGTAAAGTGTGCGCAGTTAATAGT 2999
Qy 2966 TTGCGCAAGTTGTTGCCATTTGCTACAGGATCGTGTGTCAGCTCGCTTGTGTATG 3025
Db 3000 TTGCGCAAGTTGTTGCCATTTGCTACAGGATCGTGTGTCAGCTCGCTTGTGTATG 3059

Qy 3026 GCTTCATTTACGCTCCGGTCCCAACGATCAAGCGAGTTATCATGATCCCCCATGTTGTGC 3085
Db 3060 GCTTCATTTACGCTCCGGTCCCAACGATCAAGCGAGTTATCATGATCCCCCATGTTGTGC 3119
Qy 3086 AAAAAAGCGTTAGCTTCCCTCGGTCTCCGATCGTTGTCAAGAGTAAAGTTGGCGCGAGTG 3145
Db 3120 AAAAAAGCGTTAGCTTCCCTCGGTCTCCGATCGTTGTCAAGAGTAAAGTTGGCGCGAGTG 3179
Qy 3146 TTATCACTCATGTTATGCGAGCACTGCATAAATCTCTTACTGTCTATGCCATCCGTAAGA 3205
Db 3180 TTATCACTCATGTTATGCGAGCACTGCATAAATCTCTTACTGTCTATGCCATCCGTAAGA 3239
Qy 3206 TGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCAATCTCGAATAAGTGTATGCGGCGA 3265
Db 3240 TGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCAATCTCGAATAAGTGTATGCGGCGA 3299
Qy 3266 CCGAGTTGCTCTTCCCGCGGCTCAATACCGGATAAATACCGCGCCACATAGCAGAACTTTA 3325
Db 3300 CCGAGTTGCTCTTCCCGCGGCTCAATACCGGATAAATACCGCGCCACATAGCAGAACTTTA 3359
Qy 3326 AAAGTCTCATCATTTGAAAAACGTTCTTCGGGCGGAAAACTCTCAAGGATCTTACCGCTG 3385
Db 3360 AAAGTCTCATCATTTGAAAAACGTTCTTCGGGCGGAAAACTCTCAAGGATCTTACCGCTG 3419
Qy 3386 TTGAGATCCAGTTTCGATGTAACCCACTCGTGTGCAACCAACTGTATCTTCAGCATCTTTTACT 3445
Db 3420 TTGAGATCCAGTTTCGATGTAACCCACTCGTGTGCAACCAACTGTATCTTCAGCATCTTTTACT 3479
Qy 3446 TTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGCCAAATGCCGCAAAAAGGGAATA 3505
Db 3480 TTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGCCAAATGCCGCAAAAAGGGAATA 3539
Qy 3506 AGGGGACACGGAAATGTTGAATATCATATCTCTCTTTTCAATATTTATTTGAAGCATT 3565
Db 3540 AGGGGACACGGAAATGTTGAATATCATATCTCTCTTTTCAATATTTATTTGAAGCATT 3599
Qy 3566 TATCAGGGTTATGTTCTCATGAGCGGATACATATTTGAATGTTATTTAGAAAAATAACAA 3625
Db 3600 TATCAGGGTTATGTTCTCATGAGCGGATACATATTTGAATGTTATTTAGAAAAATAACAA 3659
Qy 3626 ATAGGGTTCCCGCGACATTTCCCGGAAAAAGTCCAC 3662
Db 3660 ATAGGGTTCCCGCGACATTTCCCGGAAAAAGTCCAC 3696

RESULT 4

US-09-503-222-6
; Sequence 6, Application US/09503222
; Patent No. 6265548
; GENERAL INFORMATION:
; APPLICANT: Pavlakis, George N.
; APPLICANT: Gaitanaris, George A.
; APPLICANT: Stauber, Roland H.
; APPLICANT: Vournakis, John N.
; TITLE OF INVENTION: Mutant Aequorea victoria Fluorescent
; TITLE OF INVENTION: Proteins Having Increased Cellular Fluorescence
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/503,222
; FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA: US/08/646, 538
FILING DATE: 08/08/2005
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Kenneth A.
REGISTRATION NUMBER: 31,677
REFERENCE/DOCKET NUMBER: 015280-249000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 3699 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: - 1..3699
LOCATION: 1..3699
OTHER INFORMATION: /note= "pBSGFP"
US-09-503-222-6

Query Match 71.6%; Score 2620.6; DB 3; Length 3699;
Best Local Similarity 83.3%; Pred. No. 0;
Matches 3078; Conservative 0; Mismatches 579; Indels 40; Gaps 7;

QY	3	AAATTGTAAGCTTAATTTTGTAAATTCGCGTTAAATTTTGTAAATTCAGCTCAT	62
Db	3	AAATTGTAACGTTAATTTTGTAAATTCGCGTTAAATTTTGTAAATTCAGCTCAT	62
QY	63	TTTTTAAACCAATAGCGGAAATCGCAAAATCCCTTATAAATCAAAAGATAGACCGAGA	122
Db	63	TTTTTAAACCAATAGCGGAAATCGCAAAATCCCTTATAAATCAAAAGATAGACCGAGA	122
QY	123	TAGGGTTGAGTGTGTTCCAGTTTGGAAACAGAGTCCCACTATTAAAGAACGTGGACTCCA	182
Db	123	TAGGGTTGAGTGTGTTCCAGTTTGGAAACAGAGTCCCACTATTAAAGAACGTGGACTCCA	182
QY	183	AGTCAAAGGGCGAAACCGTCTATCAGGGCGATGGCCCACTAGTGAAACCATCACCT	242
Db	183	AGTCAAAGGGCGAAACCGTCTATCAGGGCGATGGCCCACTAGTGAAACCATCACCT	242
QY	243	AATCAAGTTTTTGGGGTTCGAGTGCCTGAAAGCACTAAATCGGAACCTTAAGGGAGCC	302
Db	243	AATCAAGTTTTTGGGGTTCGAGTGCCTGAAAGCACTAAATCGGAACCTTAAGGGAGCC	302
QY	303	CCCGATTTAGAGCTTTGACGGGAAAGCGCGCAACGTTGGCGAAGAAAGGAAAGAAAG	362
Db	303	CCCGATTTAGAGCTTTGACGGGAAAGCGCGCAACGTTGGCGAAGAAAGGAAAGAAAG	362
QY	363	CGAAAGGAGCGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCAACGTCGCGGTAAACCA	422
Db	363	CGAAAGGAGCGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCAACGTCGCGGTAAACCA	422
QY	423	CACCCGCGCGCTTAATGCGCGCTACAGGGCGGCTCCATTCGCGCATTCAGGCTGC	479
Db	423	CACCCGCGCGCTTAATGCGCGCTACAGGGCGGCTCCATTCGCGCATTCAGGCTGC	482
QY	480	GCAACTGTTGGGAAGGCGATCGGTGCGGCGCTCTTCGCTATTACGCCAGCTGGCGAAG	539
Db	483	GCAACTGTTGGGAAGGCGATCGGTGCGGCGCTCTTCGCTATTACGCCAGCTGGCGAAG	542
QY	540	GGGGATGTCGCAAGGCGATTAAGTTGGGTAAACCGCAGGGTTTTCCCGTCAACAGCTT	599
Db	543	GGGGATGTCGCAAGGCGATTAAGTTGGGTAAACCGCAGGGTTTTCCCGTCAACAGCTT	602
QY	600	GTAAACGAGCGCGCTAGCGCGCTCTTCATTCACGTTTTTGAACCCGTTGAGGAGC	659
Db	603	GTAAACGAGCGCGCTAGTTGTTGTAAT--ACGACTCACTATAGGGCGGAATTGGGTACCGG	660
QY	660	GGCAGACTCGCGGTGCAAAATGTTTAAAGCGGTGATGAGCAGATGAAGATGCTCGACA	719

Db	661	GCCCCCCTCGAGGTCGACGGTATCGATAAGCTTGATGATCCTTATTTGTATAGTTCAATC	720
QY	720	CGCTGCA---GAACACGAGCTAGATTAAACCTGAAAGATAATCATATTGTGACGTAGC	776
Db	721	CATGCCATGTGTAAATCCAGCAGCTGTTTCAAACTCAAGAAAGGACCATGTGTCTCTCTT	780
QY	777	TTAAAGATAATCATGCGTAAATTTGACGATCGGATCTGTAATACGACTCATATATAGGC	836
Db	781	TTGTTGGGATCTTTGAAAGGGCAGATTGTGTGACAGGTA-ATGGTTGTCTGGTAAAA	839
QY	837	GAAATTGGGTACCGGGCCCCCTCGAGTTCGAGGTTCGATATCGATAAGCTGATATCGAATTC	896
Db	840	GGACAGGGCCATCGCCAATTGGAGTATTTTGTGATAATGCTCTAGTTGAAACGCTTC	899
QY	897	CTGACGCGGGGGATCCACTAGTT---CTAGACGGCGGCCCAACCGGCTGGAGCTCCA	952
Db	900	CATCTTCAATGTGTCTAAATTTTGAAGTTAACTTTGATTCCTCTTTGTTGTTCTG	959
QY	953	GCTTTTGTCCCTTTAGTGAGGGTTAAATTAGATCCCATCGCTCAATTTTACGACACTAT	1012
Db	960	CCATGATGTATACATTGTGTGATTATAGTTGATTTCCAAATTTGTGTCCAAGATGTTTC	1019
QY	1013	CTTCTAGGGTTAATCTAGCTGCATCAGGATCATATCGTGGGTCTTTTTCGCGCTCAG	1072
Db	1020	CATCTTCTTTAAATCAATACCTTTTAACTCGATTCTATTAACAAGGGTATCACCTTCAA	1079
QY	1073	TCATCGCCCAAGCTGGCGCTATCTGGGCATCGGGAGGAGGAGGAGCCGCTTTCCTCGG	1132
Db	1080	ACTTGACTTCAGCAGCTGTCTGTAGTTCCCGTCACTTTTGAATAATAGTCTCTTCTCT	1139
QY	1133	CGAGTTGAAGCGGCATGGAAGAGTTTGGCGAGGATGACTGCTGCTGATTTGAGCTTGA	1192
Db	1140	GTAATAACCTTCGGGCATGGCACTTTGAAAGAGTCTGCGTTTTCATATGATCOGGT	1199
QY	1193	GGAAAAACGCTTTACATGATGATTGGGAAAGTGTGGCCATGCGACGCTTTTCGACACAGT	1252
Db	1200	ATCTTGAAGAGCATTTGAACACCAATAAGAGAAAGTAGTGACAAGTGTGGCCATGGAACAG	1259
QY	1253	TGAAGTGTGTTTCGTCAGGCCAATCGGATACCAAGTTCGTCGGGCTTTTCGCGACACAGT	1312
Db	1260	GTAGTTTTCAGTAGTGTCAAAATAAATTTAAGGGTAAAGTTTTCGCTATGTTGATCACCTT	1319
QY	1313	CGGATGTGACGCCGGAAGCGCATCAGCAACCGCAACAAATACCGCGCAGCAGCGGAACTG	1372
Db	1320	CACCTCTCCACTGACAGAAATTTGTGCCCATTAACATCACCATCTAATTAACAAGAA	1379
QY	1373	CGTGCC-----GGTGTGCAGATTAATGACAGCGGTGCGCGCTGGGATAT	1418
Db	1380	TTGGACAACTCCAGTGAAGAGTCTTCTCTTTGCTAGCCATTTCTTGGCGATCGAAT	1439
QY	1419	TACGTACGAGGAGCGGTATCTCTGGTGGATGCCGCAAGAAATGACATGGATACCCCGT	1478
Db	1440	TCCTGACGCCCGGGGATCCACTAGTTCTAGAGCGCGCCCAACCGCGTGGAGCTCCAGC	1499
QY	1479	GAGTTACCCCGCGG-----CGGCTTTGGCGTAAATCATATGCTCATAGCTGTTT	1525
Db	1500	TTTGTCTCTTTAGTGAGGGTTAATTCGAGCTTTGGGCTAATCATGTCTATAGCTGTT	1559
QY	1526	CCTGTGAAATTTGTTATCCGCTCAAAATTCACAAACATACGAGCGCGAAGCAATAAG	1585
Db	1560	CCTGTGAAATTTGTTATCCGCTCAAAATTCACAAACATACGAGCGCGAAGCAATAAG	1619
QY	1586	TGTAAGCCTGGGTGCTTAATGAGTGAAGTCAATTAATTTGCTGCTGCTCAGT	1645
Db	1620	TGTAAGCCTGGGTGCTTAATGAGTGAAGTCAATTAATTTGCTGCTGCTCAGT	1679
QY	1646	CCCGCTTTCCAGTCGGGAAACCTGTGTCAGCTGCTAATTAATGATCGGCAACCGCGG	1705
Db	1680	CCCGCTTTCCAGTCGGGAAACCTGTGTCAGCTGCTAATTAATGATCGGCAACCGCGG	1739
QY	1706	GGGAGAGCGGTTTGTGCTATTTGGGGCTCTTCCGCTTCTCGCTCACTGCTCGCTGGC	1765

Db 1740 GGGAGAGCGGTTTGGGTATTGGCGCTCTTTCGGCTTCTCGTCTACTGACTCGCTGCGC 1799
Qy 1766 TCGGTGTTTGGCTGCGGAGCGGTATCAGCTCACTCAAGAGCGGTATATCGGTTATCC 1825
Db 1800 TCGGTGTTTGGCTGCGGAGCGGTATCAGCTCACTCAAGAGCGGTATATCGGTTATCC 1859
Qy 1826 ACAGATCAGGGGATACCCAGGAAGACATGTGAGCAAAAGGCGAGCAAAAGGCGCAG 1885
Db 1860 ACAGATCAGGGGATACCCAGGAAGACATGTGAGCAAAAGGCGAGCAAAAGGCGCAG 1919
Qy 1886 AACCGTAAAAAGCGCGTGTGCGGTTTTCATAGAGCTCGCGCCCTCAGCAGCAT 1945
Db 1920 AACCGTAAAAAGCGCGTGTGCGGTTTTCATAGAGCTCGCGCCCTCAGCAGCAT 1979
Qy 1946 CACAAAAATCGACGCTCAAGTCAGAGTGGCGGAAACCCAGCAGGACTATAAGATACAG 2005
Db 1980 CACAAAAATCGACGCTCAAGTCAGAGTGGCGGAAACCCAGCAGGACTATAAGATACAG 2039
Qy 2006 GCGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTCTGTTCCGACCTCGCGCTTACCGA 2065
Db 2040 GCGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTCTGTTCCGACCTCGCGCTTACCGA 2099
Qy 2066 TACCTGTCCGCTTCTCCCTTCGGAAGCGTGGCGCTTCTCATAGCTCAGCTGTAGG 2125
Db 2100 TACCTGTCCGCTTCTCCCTTCGGAAGCGTGGCGCTTCTCATAGCTCAGCTGTAGG 2159
Qy 2126 TATCTCAGTTCGGTGTAGTTCGCTCCAGCTGGGCTGTGTGCAAGAACCCCGCTT 2185
Db 2160 TATCTCAGTTCGGTGTAGTTCGCTCCAGCTGGGCTGTGTGCAAGAACCCCGCTT 2219
Qy 2186 CAGCCCGAGCGCTGCGCTTATCCGTTAACTATCGTCTTGTAGTTCAGCCGCTAAGAC 2245
Db 2220 CAGCCCGAGCGCTGCGCTTATCCGTTAACTATCGTCTTGTAGTTCAGCCGCTAAGAC 2279
Qy 2246 GACTTATCCGCTGCGCAGCAGCTGTGTACAGGATTTAGCAGCGAGGTATGTAGG 2305
Db 2280 GACTTATCCGCTGCGCAGCAGCTGTGTACAGGATTTAGCAGCGAGGTATGTAGG 2339
Qy 2306 GGTGTCTACAGAGTTCCTTGAAGTGGTGGCTTAACTACGCTACACTAGAGGACAGTATTT 2365
Db 2340 GGTGTCTACAGAGTTCCTTGAAGTGGTGGCTTAACTACGCTACACTAGAGGACAGTATTT 2399
Qy 2366 GGTATCTGCGCTCTGCTGAAGCAGTTACTTCGGAAGAGAGTGGTAGCTCTTGATCC 2425
Db 2400 GGTATCTGCGCTCTGCTGAAGCAGTTACTTCGGAAGAGAGTGGTAGCTCTTGATCC 2459
Qy 2426 GGCBAACCAACACCGCTGTTAGCGGTGTTTTTTTGTGCAAGCAGAGATTACCGC 2485
Db 2460 GGCBAACCAACACCGCTGTTAGCGGTGTTTTTTTGTGCAAGCAGAGATTACCGC 2519
Qy 2486 AGAAAAAAGGATCTCAAGAAGATCTTTGATCTTTTCTACGGGCTCTGACGCTCAGTGG 2545
Db 2520 AGAAAAAAGGATCTCAAGAAGATCTTTGATCTTTTCTACGGGCTCTGACGCTCAGTGG 2579
Qy 2546 AACGAAAACTCAGTTAAGGATTTTGGTCATGAGATTATCAAAAAGGATCTTACCTAG 2605
Db 2580 AACGAAAACTCAGTTAAGGATTTTGGTCATGAGATTATCAAAAAGGATCTTACCTAG 2639
Qy 2606 ATCTTTTAAATTAATTAAGTATTAATCAATCAATTAAGTATATATGATTAAGTTCGG 2665
Db 2640 ATCTTTTAAATTAATTAAGTATTAATCAATCAATTAAGTATATATGATTAAGTTCGG 2699
Qy 2666 TCTGACAGTTACCAATGCTTAATCAGTGAAGGACCTATCTCAGCGATCTGTCTATTTCGT 2725
Db 2700 TCTGACAGTTACCAATGCTTAATCAGTGAAGGACCTATCTCAGCGATCTGTCTATTTCGT 2759
Qy 2726 TCATCCATAGTTCCTGACTCCCGCTGTTAGATAACTACGATACGGAGGCTTACCA 2785
Db 2760 TCATCCATAGTTCCTGACTCCCGCTGTTAGATAACTACGATACGGAGGCTTACCA 2819
Qy 2786 TCTGGCCCGAGTCTGCAATGATACCGGAGACCCAGCTCAGCTCCAGTTATCA 2845
Db 2820 TCTGGCCCGAGTCTGCAATGATACCGGAGACCCAGCTCAGCTCCAGTTATCA 2879

Qy 2846 GCAATAAACACGACGCGGAAAGGGCCGAGCGCAGAGTAAGTGTCTCTCAAACTTTTATCGGC 2905
Db 2880 GCAATAAACACGACGCGGAAAGGGCCGAGCGCAGAGTAAGTGTCTCTCAAACTTTTATCGGC 2939
Qy 2906 TCCATCCAGTCTATTAAATTTGTCGGGAAGCTTAGAGTAAGTAGTTCGCCAGTTAATAGT 2965
Db 2940 TCCATCCAGTCTATTAAATTTGTCGGGAAGCTTAGAGTAAGTAGTTCGCCAGTTAATAGT 2999
Qy 2966 TTGCGCAACGTTGTTGCGCAATTCGTACAGGCATCGTGGTGTACAGCTCGTGTGTTGATG 3025
Db 3000 TTGCGCAACGTTGTTGCGCAATTCGTACAGGCATCGTGGTGTACAGCTCGTGTGTTGATG 3059
Qy 3026 GCTTCAATTCAGCTCCGCTTCCCAACGATCAAGCGAGTTACATGATCCGCCATGTTGTGC 3085
Db 3060 GCTTCAATTCAGCTCCGCTTCCCAACGATCAAGCGAGTTACATGATCCGCCATGTTGTGC 3119
Qy 3086 AAAAAAGCGTTAGCTCTTCGCTCCTCGATCGTTGTCAAGAAAGTAAGTTGGCGCAGTG 3145
Db 3120 AAAAAAGCGTTAGCTCTTCGCTCCTCGATCGTTGTCAAGAAAGTAAGTTGGCGCAGTG 3179
Qy 3146 TTATCACTCATGTTATGCGACACTGCAATAATTTCTTTACTGTCAATGCCATCCGTAAGA 3205
Db 3180 TTATCACTCATGTTATGCGACACTGCAATAATTTCTTTACTGTCAATGCCATCCGTAAGA 3239
Qy 3206 TCGTTTTCTGACTGCTGAGTACTCAACCAAGTCAATTCGTGAGAAATAGTGTATGCGCGA 3265
Db 3240 TCGTTTTCTGACTGCTGAGTACTCAACCAAGTCAATTCGTGAGAAATAGTGTATGCGCGA 3299
Qy 3266 CCGAGTTGCTCTTCCCGCGCTCAATACGGGATATACCGCCACATAGCAGACTTTA 3325
Db 3300 CCGAGTTGCTCTTCCCGCGCTCAATACGGGATATACCGCCACATAGCAGACTTTA 3359
Qy 3326 AAAGTCTCATCATTTGGAAAAAGTTCTTCGGGCGGAAAACTCTCAAGGATCTTACCGCTG 3385
Db 3360 AAAGTCTCATCATTTGGAAAAAGTTCTTCGGGCGGAAAACTCTCAAGGATCTTACCGCTG 3419
Qy 3386 TTGAGATCAGTTGAGTAAACCCACTCGTGAACCCAACTGATCTTCAGATCTTTTACT 3445
Db 3420 TTGAGATCAGTTGAGTAAACCCACTCGTGAACCCAACTGATCTTCAGATCTTTTACT 3479
Qy 3446 TTCCAGCGTTTCTGGGTGAGCAAAACAGGAGCAAAATCCCGCAAAAGGGAATA 3505
Db 3480 TTCCAGCGTTTCTGGGTGAGCAAAACAGGAGCAAAATCCCGCAAAAGGGAATA 3539
Qy 3506 AGGGGACACGCAAAATGTTGAATATCTATCTCTCTCTTTTCAATATTATTCAGCAT 3565
Db 3540 AGGGGACACGCAAAATGTTGAATATCTATCTCTCTCTTTTCAATATTATTCAGCAT 3599
Qy 3566 TATCAGGTTATGTTCTCATGAGCGGATACATATTTGAATGTTATTTAGAAAAATAACAA 3625
Db 3600 TATCAGGTTATGTTCTCATGAGCGGATACATATTTGAATGTTATTTAGAAAAATAACAA 3659
Qy 3626 ATAGGGTTCCCGGACATTTCCCGAAAAAGTCCAC 3662
Db 3660 ATAGGGTTCCCGGACATTTCCCGAAAAAGTCCAC 3696

RESULT 5

US-09-890-848-16
; Sequence 16, Application US/09890848
; Patent No. 6573053
; GENERAL INFORMATION:
; APPLICANT: Firth, Greg
; APPLICANT: Odedra, Rajesh Muru
; TITLE OF INVENTION: Analysis Method
; FILE REFERENCE: PA9903
; CURRENT APPLICATION NUMBER: US/09/890,848
; CURRENT FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: PCT/GB00/00346
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: EP 99300873.9
; PRIOR FILING DATE: 1999-02-05

```
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 5277
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: plasmid
US-09-890-848-16

Query Match      67.4%; Score 2469.4; DB 4; Length 5277;
Best Local Similarity 82.3%; Pred. No. 0;
Matches 3014; Conservative 0; Mismatches 436; Indels 213; Gaps 7;

QY 3 AAATTGTAAGCGTTAAATTTTGTAAAAATTCGCGTTAAATTTTGTAAAAATCAGCTCAT 62
DB 1239 AAATTGTAACGTTAAATTTTGTAAAAATTCGCGTTAAATTTTGTAAAAATCAGCTCAT 1298
QY 63 TTTTAAACCAATPAGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAAATAGACCGAGA 122
DB 1299 TTTTAAACCAATPAGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAAATAGACCGAGA 1358
QY 123 TAGGTTGAGTGTGTTCCAGTTTGGAAACAAGTCCACTATTAAGAAACGTGGACTCCA 182
DB 1359 TAGGTTGAGTGTGTTCCAGTTTGGAAACAAGTCCACTATTAAGAAACGTGGACTCCA 1418
QY 183 AGCTCAAAAGGGCGAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCT 242
DB 1419 AGCTCAAAAGGGCGAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCT 1478
QY 243 AATCAAGTTTTTGGGTCGAGGTGCGGTAAAGCACTAAATCGGAACCTTAAGGGAGCC 302
DB 1479 AATCAAGTTTTTGGGTCGAGGTGCGGTAAAGCACTAAATCGGAACCTTAAGGGAGCC 1538
QY 303 CCCGATTTAGACTTGAACGGGAAACCGCGCAACGTTGGCGAAGAAAGGAGGAAGAAG 362
DB 1539 CCCGATTTAGACTTGAACGGGAAACCGCGCAACGTTGGCGAAGAAAGGAGGAAGAAG 1598
QY 363 CGAAAGGACGGCGCTAGGGCGCTGGCAAGTGTAGCGGTACCGTGGCGGTAAACCA 422
DB 1599 CGAAAGGACGGCGCTAGGGCGCTGGCAAGTGTAGCGGTACCGTGGCGGTAAACCA 1658
QY 423 CACCCGCGCGCTTAATGCGCGCTACAGGGCGCGTC---CAATTGCGCAATTCAGGCTGC 479
DB 1659 CACCCGCGCGCTTAATGCGCGCTACAGGGCGCGTC---CAATTGCGCAATTCAGGCTGC 1718
QY 480 GCAACTGTTGGGAAGGGCGATCGGTGCGGGCGCTCTTCGCTATTACGCCAGCTGGCGAAG 539
DB 1719 GCAACTGTTGGGAAGGGCGATCGGTGCGGGCGCTCTTCGCTATTACGCCAGCTGGCGAAG 1778
QY 540 GGGGATGTCTGCAAGGCGATTAAGTTGGGTAAACGCCAGGGTTTTCCAGTCACCACTGTT 599
DB 1779 GGGGATGTCTGCAAGGCGATTAAGTTGGGTAAACGCCAGGGTTTTCCAGTCACCACTGTT 1838
QY 600 GTAAACGACGCGCAGTGAAGCGCGCTGTTCAATTCAAGTTTTTGAACCCGCTGGAGGACG 659
DB 1839 GTAAACGACGCGCAGTGAAGTTGTAATACGACTCACTATAGGCGCAATTGGAGCTCGAGC 1898
QY 660 GGCAGACTCGGGTCCAAATGTGTTTTACGGGTGATGGAGCAGATGAAGTGTCTGACA 719
DB 1899 AATAAAGATTCTACAATACTAGCTTTTATGGTTTATGAAGAGGAAATAATTGGCAGTAACT 1958
QY 720 CGCTCAGAACACGACGCTAGATTAACCTAGAAAGATAATCATATTGTGACGTACGTTA 779
DB 1959 GGCCCCACAACCC-----TTCAATGAACGAATCAAAATTAACAACCATAGGATGATA 2010
QY 780 AAGATAATCATGCGTAAATTAAGCGCATGGGATCTGTAATACGACTCACTATAGGGCGAA 839
DB 2011 ATGCGATTAGTTTTTATGCTTATTTCTGGGGTAAATTAATCAAGCAAGCGAT---GATT 2066
QY 840 TTGGGTACGGGCCCCCTCGAGGTTCGACGGTATCGATAGCTTGATATCGAATTCCTG 899
```



```
; NAME/KEY: misc_feature
; LOCATION: ().T()
; OTHER INFORMATION: Synthetic sequence
US-09-402-266B-21

Query Match      66.9%; Score 2448.8; DB 4; Length 3956;
Best Local Similarity 79.4%; Pred. No. 0;
Matches 3136; Conservative 0; Mismatches 522; Indels 294; Gaps 8;

QY 5 ATTGTAAAGCGTTAAATATTTTGTAAATTCGGTAAATTTTGTAAATCAGCTCATTT 64
D 1 ATTGTAAAGCGTTAAATATTTTGTAAATTCGGTAAATTTTGTAAATCAGCTCATTT 60
QY 65 TTTTAAACCAATAGCGCGAAATCGGCAAAATCCCTTATAAATCAAAAGATAGACCGAGATA 124
D 61 TTTTAAACCAATAGCGCGAAATCGGCAAAATCCCTTATAAATCAAAAGATAGACCGAGATA 120
QY 125 GGGTTGAGTGTGTTTCCAGATTTTGGAAACAAGATGCCACTATTAAGAAAGCGTGAATCCCAAC 184
D 121 GGGTTGAGTGTGTTTCCAGTGTGGAACAAGATGCCACTATTAAGAAAGCGTGAATCCCAAC 180
QY 185 GTCAAGGGCGGAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCCTAA 244
D 181 GTCAAGGGCGGAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCCTAA 240
QY 245 TCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAAACCTTAAAGGGAGCCCC 304
D 241 TCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAAACCTTAAAGGGAGCCCC 300
QY 305 CGATTTAGAGCTTGACGGGGAAGCGCGGAACGTTGGGAGAGAAAGGGAAGAAAGCG 364
D 301 CGATTTAGAGCTTGACGGGGAAGCGCGGAACGTTGGGAGAGAAAGGGAAGAAAGCG 360
QY 365 AAAGGAGCGGCGCTAGGGCGTGGCAAGTGTAGCGGTCAAGCTGCGGTGAACCCACCA 424
D 361 AAAGGAGCGGCGCTAGGGCGTGGCAAGTGTAGCGGTCAAGCTGCGGTGAACCCACCA 420
QY 425 CCGCGCGCTTAATGCGCGCTTACAGGGCGGCTGCCATTTCGCCATTACAGGCTGCGCAAC 484
D 421 CCGCGCGCTTAATGCGCGCTTACAGGGCGGCTGCCATTTCGCCATTACAGGCTGCGCAAC 480
QY 485 TGTGGGAAGGCGATCGGTGCGGCGCTCTTCGCTATTACGCCAGCTGCGGAAAGGGGA 544
D 481 TGTGGGAAGGCGATCGGTGCGGCGCTCTTCGCTATTACGCCAGCTGCGGAAAGGGGA 540
QY 545 TGTGCTGCAAGCGGATTAAGTTGGGTAAAGCAGGGTTTTCCAGTCAAGAGTTGTAA 604
D 541 TGTGCTGCAAGCGGATTAAGTTGGGTAAAGCAGGGTTTTTCAGTCAAGAGTTGTAA 600
QY 605 ACGAGCGGCGATGAGGCGGCTCGTTTCAATTCACGTTTTTGAACCGGTGGAGGACGGGAG 664
D 601 ACGAGCGGCGATGAGGCGGCTCGTTTCAATTCACGTTTATAGGGCGAATGGAGCTCCACGC 660
QY 665 ACTCGCGGTGCAAAATGTGTTTTACAGCGGTGATGGAGC----- 701
D 661 GGTGGCGCTCTAGAACTAGTGGATCGGCGCA TGGAGGCCACACCGCATAGATCGGCAAGT 720
QY 702 ----- 701
D 721 GCACAAACAATACTTAAATAAATACTACTCAGTAATAACCTATTTCTTAGCAATTTTTCAG 780
QY 702 -----AGATGAAGATGCTCGACACGCTGAGAACACGACGAGCTAGATT 743
D 781 GAAATTTGCTATTTTGTAGAGCTTTTACACCATTTGTCTCCACCTCCGCTTACATC 840
QY 744 AACCCCTAGAAAGATATCATATTTGTGACGTA CGTTAAAGATATCATCGCT----- 794
D 841 AACACCAATAACGCCAATTTAA TCTAAGCGCATCAACCAATTTTCTGGGCTCAGTCCACC 900
QY 795 -----AAAATTGACGATGGGATCTGTAAATACGACTCACTATATAGGGGGAAT--- 841
D 901 AGCTAACATAAAATGTAAGCTTTTCGGGGCTCTCTGCTTCCAAACCCAGTCAGAAATCGA 960
```

```
QY 842 ----- 841
D 961 GTTCCAATCAAAAGTTCACTGTCTCCACCTTGCTTGAATCAAAACAGGGAATAAACGA 1020
QY 842 -----GGGTACCGGGCCCCCTCGAGGTCGACGGTATCGATAAGCTTGTATTCGATTC 897
D 1021 ATGAGGTTTCTGTGAAGCTGCACCTAGTAGTATGTTGCACTCTTTTGGAAATACGAGTCT 1080
QY 898 TGCACGCCGGGGATCCACTAGTTCTTAGAGGGCGCCACCGCGTGGAGCTTCCA----- 952
D 1081 TTTAATAACTGCAAAACGAGGAACCTCTTGTATTTCTTGCACAGACTCATCTCCATCGAG 1140
QY 953 --GCTTTTGTCCCTTTAGTGAGGGTTAATTAGATCCCATGCGTCAAAATTTTACGAGACT 1010
D 1141 TTGGACGATATCAATGCCGTAATCATTTGACAGAGCCAAACATCTCCTCTAGGTTGATT 1200
QY 1011 ATCTTTT----- 1017
D 1201 ACGAAACACGCCAACCAAGTATTTCCGAGTGCTTGAACCTATTTTATATGCTTTTACAAG 1260
QY 1018 --TAGGGTTAATCTAGCTGCATCAGGATCATATCGTGGGTCTTTTTCGGCTCAGTC 1074
D 1261 ACTTGAATTTTCTTGTGAATAACCGGTCAATGTTCTCTTTCTATTGGGCACACATAT 1320
QY 1075 ATGCGCCCAAGCTGGCGCTATCTG-----GGCATCGGGAGGAAGAA 1115
D 1321 AATACCCAGCAAGTCAGCATCGGAATCTAGAGCACATTTCTGGGCTCTGTCTCTGCAA 1380
QY 1116 GCCGTGCTTTTCCGCGAGGTTGAAGGGCATGGAAGAGTTTGGCGAGGATGACTGC 1175
D 1381 GCCGCAAACTTTTCAACCAATGGACCAAACTACTGTGAATTTAATAACAGACATACTCA 1440
QY 1176 TGCTCATTTGAGCTTGTAG-----CGAAAAAGCAGCTTTTACCATGATGATTCGG----- 1224
D 1441 AGTCCCTTTGTGCTTAAATCACGTATCTCACGTGCTCAATAGTACCAATGCCCTCC 1500
QY 1225 -----AAGGTGGCCATGACGCTTTTAAACGGTGAACCTGTTGCTTCAAGC 1270
D 1501 CTCCTTGGCCCTCTCCTTTTCTTTTTCGACCGAA TTAATTTCTTAATCGSCAAAAAGAA 1560
QY 1271 CACTGGGATACAGTTCTGTCGGCTTTTTCGGACACAGTTCCGGATGGTCAGCCGNA 1330
D 1561 AAGCTCCGGATCAAGATTGTACGTAAGGTGAACAAGCTATTTTCAATAAAGAAATATCTTC 1620
QY 1331 GCGCATCAGCAACCCGAACAATACCGCGACAGCGCGAACTGCGGTGCGGTGTCAGAT 1390
D 1621 CACTACTGCATCTGCGCTCATACTGCAAGTAGTACACATATATACGATGCTGTCTATT 1680
QY 1391 TAAATGACAGCGTGGCGCTGGGATATTACGTACAGGAGGACGGGTATCTCTGGCTGGAT 1450
D 1681 AAATGCGGCGCTGCAGGAATTCGATATCAAGCTTATCGATACCGCTCGACCTCGAGGGG 1740
QY 1451 GCGCAGAAATGGACATGGATACCCCGTAGTTACCGCGGGCGGCTTGGCGTAAATCA 1510
D 1741 GCGCGGATACCCAGCTTTTGTTCCTTTAGTAGGGTTAAATTGCGCGCTTGGCGTAAATCA 1800
QY 1511 TGGTCATAGCTGTTCTCTGTGTAATTTGTTATCGCTCACAAATTCACACAAACATACGA 1570
D 1801 TGGTCATAGCTGTTCTCTGTGTAATTTGTTATCGCTCACAAATTCACACAAACATACGA 1860
QY 1571 GCGGAAGCATAAAGTGTAAAGCTGGGGTGCCTTAATGAGTGAGTAACTCACTCAATTAAT 1630
D 1861 GCGGAAGCATAAAGTGTAAAGCTGGGGTGCCTTAATGAGTGAGTAACTCACTCAATTAAT 1920
QY 1631 GGGTTGCGCTCACTGCGGCTTTCCAGTCGGGAACCTGTGTCGAGCTGCATTAATGA 1690
D 1921 GGGTTGCGCTCACTGCGGCTTTCCAGTCGGGAACCTGTGTCGAGCTGCATTAATGA 1980
QY 1691 ATCGGCCAAAGCGGGGAGAGCGGTTTGGCTATTTGGCGCTCTTTCGCTCTCTCGCTC 1750
D 1981 ATCGGCCAAAGCGGGGAGAGCGGTTTGGCTATTTGGCGCTCTTTCGCTCTCTCGCTC 2040
QY 1751 ACTGACTCGCTCGCTCGGTGTTTCGGCTGCGGCGAGCGGTTATCAGCTCACTCAAGGCG 1810
```


[illegible]

QY 1485 CCGGCGGGCGCGCTTGGCGTAATCAATGATCATAGCTGTTTCTGTGTGAAATTTGTTATC 1544
DB 1921 GGTAAATTTGGCGCTTGGCGTAATCATGTGTCATAGCTGTTTCTGTGTGAAATTTGTTATC 1980
QY 1545 CGCTCACAATTCACACACATACGAGCGGAGAGCATAAAGTGTAAAGCTGGGGTGCCCT 1604
DB 1981 CGCTCACAATTCACACACATACGAGCGGAGAGCATAAAGTGTAAAGCTGGGGTGCCCT 2040
QY 1605 AATGAGTGAGCTAACTCACATTAATTTGCGTTGCGCTCACTGCCCGCTTTCCAGTCCGGAA 1664
DB 2041 AATGAGTGAGCTAACTCACATTAATTTGCGTTGCGCTCACTGCCCGCTTTCCAGTCCGGAA 2100
QY 1665 ACTGTGCTGCGAGCTGCAATTAATGAATCGGCCAACCGCGGGGAGAGCGGTTTCGGTA 1724
DB 2101 ACTGTGCTGCGAGCTGCAATTAATGAATCGGCCAACCGCGGGGAGAGCGGTTTCGGTA 2160
QY 1725 TTGGGCGCTCTTCCGCTTCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCCGCTCGGC 1784
DB 2161 TTGGGCGCTCTTCCGCTTCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCCGCTCGGC 2220
QY 1785 GAGCGGTATCAGCTCACTCAAAAGCGGTAAATACGTTATCCACAGAAATCAGGGGATAACG 1844
DB 2221 GAGCGGTATCAGCTCACTCAAAAGCGGTAAATACGTTATCCACAGAAATCAGGGGATAACG 2280
QY 1845 CAGGAAGAACATGTGAGCAAAAGGCGAGCAAAAGGCGAGAACCGTAAAGGCGCGGT 1904
DB 2281 CAGGAAGAACATGTGAGCAAAAGGCGAGCAAAAGGCGAGAACCGTAAAGGCGCGGT 2340
QY 1905 TGCCTGCGCTTTTCCATAGCTCCGCCCCCTGACGAGCATCAAAAATCGAGCTCAA 1964
DB 2341 TGCCTGCGCTTTTCCATAGCTCCGCCCCCTGACGAGCATCAAAAATCGAGCTCAA 2400
QY 1965 GTCAGAGGTGGCGAAACCCGACAGACTATAAGATACACAGGCGTTTCCCTCGGAAGCT 2024
DB 2401 GTCAGAGGTGGCGAAACCCGACAGACTATAAGATACACAGGCGTTTCCCTCGGAAGCT 2460
QY 2025 CCTCGTGGCTCTCTGTTTCCAGCTCGCGCTTACCGGATACCTGTCCGCTTTCTCC 2084
DB 2461 CCTCGTGGCTCTCTGTTTCCAGCTCGCGCTTACCGGATACCTGTCCGCTTTCTCC 2520
QY 2085 CTTGCGGAAGCGTGGCGCTTCTCATAGCTCAGCTGATAGTTATCTCAGTTGGGTAGG 2144
DB 2521 CTTGCGGAAGCGTGGCGCTTCTCATAGCTCAGCTGATAGTTATCTCAGTTGGGTAGG 2580
QY 2145 TCGTTCGCTCCAGCTGGGCTGTGTCAGCAACCCCGCTTACGCCGACCGCTGCGCT 2204
DB 2581 TCGTTCGCTCCAGCTGGGCTGTGTCAGCAACCCCGCTTACGCCGACCGCTGCGCT 2640
QY 2205 TATCCGGTAACTATCGTCTTGAAGTCCAAACCCCGTAAAGACAGACTTATCGCCACTGGCAG 2264
DB 2641 TATCCGGTAACTATCGTCTTGAAGTCCAAACCCCGTAAAGACAGACTTATCGCCACTGGCAG 2700
QY 2265 CAGCCTATGTAACAGATTTAGCAGAGCGAGGTATGTAGCGGTGTACAGAGTTCTTGA 2324
DB 2701 CAGCCTATGTAACAGATTTAGCAGAGCGAGGTATGTAGCGGTGTACAGAGTTCTTGA 2760
QY 2325 AGTGTGGCTAACTACGCTACACTAGAGAGCAGTATTTGTATCTGCGCTCTGCTGA 2384
DB 2761 AGTGTGGCTAACTACGCTACACTAGAGAGCAGTATTTGTATCTGCGCTCTGCTGA 2820
QY 2385 AGCAGTTACCTTCGGAAAAAGTGTGTAGCTCTTGTATCGCGCAACCAACCCCGCTG 2444
DB 2821 AGCAGTTACCTTCGGAAAAAGTGTGTAGCTCTTGTATCGCGCAACCAACCCCGCTG 2880
QY 2445 GTAGCGGTGTTTTTTTTTGAAGCAGAGATTAACGCGAGAAAAAGAGTCTCAAG 2504
DB 2881 GTAGCGGTGTTTTTTTTTGAAGCAGAGATTAACGCGAGAAAAAGAGTCTCAAG 2940
QY 2505 AAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGGAGCAAACTCAGCTTAAG 2564
DB 2941 AAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGGAGCAAACTCAGCTTAAG 3000

QY 2565 GGATTTTGTGTCATGAGATTATCAAAAAGGATCTTCACTAGATCCTTTTAAATTAATAAT 2624
DB 3001 GGATTTTGTGTCATGAGATTATCAAAAAGGATCTTCACTAGATCCTTTTAAATTAATAAT 3060
QY 2625 GAAGTTTTAAATCAATCTAAAGTATATATAGTAAACTTTGGTCTGACAGTTACCAATGCT 2684
DB 3061 GAAGTTTTAAATCAATCTAAAGTATATATAGTAAACTTTGGTCTGACAGTTACCAATGCT 3120
QY 2685 TAATCAGTGAGCGACTATCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTTGCTGAC 2744
DB 3121 TAATCAGTGAGCGACTATCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTTGCTGAC 3180
QY 2745 TCCCGCTGCTGTAGATAACTACGATACGAGGGGCTTACCATCTGCGCCAGTGTGCA 2804
DB 3181 TCCCGCTGCTGTAGATAACTACGATACGAGGGGCTTACCATCTGCGCCAGTGTGCA 3240
QY 2805 TGATACCGCGAGACCCACGCTCAACGCTCAGATTTATCAGCAATAAACACAGCGAGCG 2864
DB 3241 TGATACCGCGAGACCCACGCTCAACGCTCAGATTTATCAGCAATAAACACAGCGAGCG 3300
QY 2865 GAAGGCGCGAGCGAGAGTGTCTCTGCAACTTTTATCCGCTCCATCCAGTCTATTAAT 2924
DB 3301 GAAGGCGCGAGCGAGAGTGTCTCTGCAACTTTTATCCGCTCCATCCAGTCTATTAAT 3360
QY 2925 GTTTCGCGGAACTAGAGTAAGTGTGCGCAGTTAATAGTTTTCGCAACGTTGTTGCCA 2984
DB 3361 GTTTCGCGGAACTAGAGTAAGTGTGCGCAGTTAATAGTTTTCGCAACGTTGTTGCCA 3420
QY 2985 TTGCTACAGGCATCGTGGTGTACAGCTCGTCTGTTGGTATGCTTCAATCAGCTCCGGTT 3044
DB 3421 TTGCTACAGGCATCGTGGTGTACAGCTCGTCTGTTGGTATGCTTCAATCAGCTCCGGTT 3480
QY 3045 CCCAACGATCAAGCGGAGTTACATGATCCCCATGTTGTGCAAAAAGCGTTAGTCTCT 3104
DB 3481 CCCAACGATCAAGCGGAGTTACATGATCCCCATGTTGTGCAAAAAGCGTTAGTCTCT 3540
QY 3105 TCGGCTCTCCGATCGTGTGTCAGAGTAAGTGTGCGCGCAGTGTATCACTCATGTTATGG 3164
DB 3541 TCGGCTCTCCGATCGTGTGTCAGAGTAAGTGTGCGCGCAGTGTATCACTCATGTTATGG 3600
QY 3165 CAGCACTGCATTAATCTCTTACTGTCTATGCCATCGTAAGATGCTTTCTGTGACTGGTG 3224
DB 3601 CAGCACTGCATTAATCTCTTACTGTCTATGCCATCGTAAGATGCTTTCTGTGACTGGTG 3660
QY 3225 AGTACTCAACCAAGTCAITCTGAGAAATAGTGTATGCGGCGACCGAGTGTCTCTTCCCGG 3284
DB 3661 AGTACTCAACCAAGTCAITCTGAGAAATAGTGTATGCGGCGACCGAGTGTCTCTTCCCGG 3720
QY 3285 CGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAGTGCTCATCATTTGGAA 3344
DB 3721 CGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAGTGCTCATCATTTGGAA 3780
QY 3345 AACGTTCTTTCGGGGCGAAACTCTCAAGGATCTTACCGCTGTGAGATCCAGTTCGATGT 3404
DB 3781 AACGTTCTTTCGGGGCGAAACTCTCAAGGATCTTACCGCTGTGAGATCCAGTTCGATGT 3840
QY 3405 AACCCACTCGTGCACCCAACTGATCTTTCAGCATCTTTTTCACACAGCGTTTCTGGGT 3464
DB 3841 AACCCACTCGTGCACCCAACTGATCTTTCAGCATCTTTTTCACACAGCGTTTCTGGGT 3900
QY 3465 GAGCAAAAACAGGAGGCAAAAATGCGCAAAAAGGAAATAAGGGCGACACGGAAATGTT 3524
DB 3901 GAGCAAAAACAGGAGGCAAAAATGCGCAAAAAGGAAATAAGGGCGACACGGAAATGTT 3960
QY 3525 GAATCTCATCTCTCTCTTTTTCATATTTAGAGCATTTATCAGGGTTATTTGCTCA 3584
DB 3961 GAATCTCATCTCTCTCTTTTTCATATTTAGAGCATTTATCAGGGTTATTTGCTCA 4020
QY 3585 TGAGCGGATACATATTTTGAATGTATTTAGAAAAATAAAGAGGGGTTCCGCGCACAT 3644
DB 4021 TGAGCGGATACATATTTTGAATGTATTTAGAAAAATAAAGAGGGGTTCCGCGCACAT 4080
QY 3645 TTCCCGGAAAAAGTGCCAC 3662

Db 4081 TTCCCGGAAAGTGCCAC 4098
|||||
RESULT 8
US-09-402-266B-18
; Sequence 18, Application US/09402266B
; Patent No. 6537767
; GENERAL INFORMATION:
; APPLICANT: HINNEN, Albert
; APPLICANT: HEGEMANN, Johannes
; APPLICANT: MUNDER, Thomas
; APPLICANT: SCHUSTER, Tilmer
; APPLICANT: FELDWMANN, Horst
; APPLICANT: KRAMER, Wilfried
; APPLICANT: ZIMMERMANN, Friedrich
; APPLICANT: ENTIAN, Karl-Dieter
; TITLE OF INVENTION: Process for Screening Antimycotically Active Substances
; FILE REFERENCE: 38005-0094
; CURRENT APPLICATION NUMBER: US/09/402,266B
; PRIOR FILING DATE: 1999-10-08
; PRIOR APPLICATION NUMBER: PCT/EP98/01904
; PRIOR FILING DATE: 1998-10-08
; PRIOR APPLICATION NUMBER: DE 19713572.2
; PRIOR FILING DATE: 1997-04-02
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 18
; LENGTH: 4088
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1..T)
; OTHER INFORMATION: Synthetic sequence
US-09-402-266B-18
Query Match 62.9%; Score 2302; DB 4; Length 4088;
Best Local Similarity 76.1%; Pred. No. 0;
Matches 3108; Conservative 0; Mismatches 550; Indels 426; Gaps 5;
QY 5 ATTGTAAGCGTTAAATTTTGTAAATTTTGGGTTAAATTTTGGTTAAATCAGCTCATTT 64
Db 1 ATTGTAAGCGTTAAATTTTGTAAATTTTGGGTTAAATTTTGTAAATGACCTCATTT 60
QY 65 TTAAACCAATAGCGGAAATCGGCAAAATCCCTTATAAATCAAAAGATAGACCCAGATA 124
Db 61 TTAAACCAATAGCGGAAATCGGCAAAATCCCTTATAAATCAAAAGATAGACCCAGATA 120
QY 125 GGGTTGAGTGTGTTCCAGTTTGGAAACAGAGTCCCACTATTAAAGAACTGGACTCCCAAC 184
Db 121 GGGTTGAGTGTGTTCCAGTTTGGAAACAGAGTCCCACTATTAAAGAACTGGACTCCCAAC 180
QY 185 GTCAAGGGCGGAAACCGTCTATCAGGGCGATGCGCCATACAGTGAACCATCATCCCTAA 244
Db 181 GTCAAGGGCGGAAACCGTCTATCAGGGCGATGCGCCATACAGTGAACCATCATCCCTAA 240
QY 245 TCAAGTTTGGGGTCCAGGTCGCGTAAAGCACTAAATCGGAACCCCTAAAGGGAGCCCC 304
Db 241 TCAAGTTTGGGGTCCAGGTCGCGTAAAGCACTAAATCGGAACCCCTAAAGGGAGCCCC 300
QY 305 CGATTTAGAGCTTGACGGGGAAGCGCGCAACGTTGGGAGAGAAAGGGAAGAAAGCG 364
Db 301 CGATTTAGAGCTTGACGGGGAAGCGCGCAACGTTGGGAGAGAAAGGGAAGAAAGCG 360
QY 365 AAAGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCAAGTGGCGGTAAACCAACA 424
Db 361 AAAGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCAAGTGGCGGTAAACCAACA 420
QY 425 CCCGCGGGCTTAAATGCGCGCTACAGGCGGCTCCCATTTGCCATTGAGGCTGCGCAAC 484
Db 421 CCCGCGGGCTTAAATGCGCGCTACAGGCGGCTCCCATTTGCCATTGAGGCTGCGCAAC 480

QY 485 TGTTCGGAAGGGCGATCGGTGCGGCGCTCTTCGCTATTATACGCCAGCTCGCGAAAGGGGA 544
Db 481 TGTTCGGAAGGGCGATCGGTGCGGCGCTCTTCGCTATTATACGCCAGCTCGCGAAAGGGGA 540
QY 545 TGTGCTGCAAGGGCGATTAAAGTTGGTAAAGCGGAGGTTTCCAGTACAGGCTTGTATAA 604
Db 541 TGTGCTGCAAGGGCGATTAAAGTTGGTAAAGCGGAGGTTTCCAGTACAGGCTTGTATAA 600
QY 605 ACGACGGCAGTGAGCGCGCTCGTTTCATTTCAGTTTGTGAACCGCTGAGGACGGGCG 664
Db 601 ACGACGGCAGTGAGCGCGCTCGTTTCATTTCAGTTTCATATAGGGGGAATTGGAGCTCCACCG 660
QY 665 ACTCGCGTGCATAATGTGTTTACAGCGTGTAGGAG----- 700
Db 661 GGTGGCGCTCTAGAACTAGTGGATCCGGCCATGGAGGCTCGTTTCAGAAATGACAGTATA 720
QY 701 --CAGATGAAGTGTGACAGCGCTGACAGACAGCAGCTAGATTAAACCTAGAAAGATA 758
Db 721 GAATGATGCAATTACCTTGTCTATCTTTCAGTATACATCTGTTGATATACATCTTCTGACA 780
QY 759 ATCATATTGTGACGTACGTTAAAGATAATCATGCGTAAATTTGACGCATGGGATCTGTAA 818
Db 781 TTATAGGTATACATATATACATGATATATATATATATATATATATATATATATATATATAT 840
QY 819 TACGACTCATATAGGGCGAATTGGGTACCGGGCCCCCTCGAGGTGACGCTATCGAT 878
Db 841 GTGTCACATACATAGAACACCTTTGGTGGAGGAAACATCGTTGGTACCATTTGGGCGAGT 900
QY 879 AAGCTTGATATCGAATTCCTGACGCGCGGGGATCCTAGTCTTAGAG----- 927
Db 901 GGCCTTCTTATGGCAACCGCAAGAGCCTTGAACCACTCTCACACGCGTGAATGATCATTT 960
QY 928 -----CGGCGGCCACCGCGTGGAGCTCCAGCTTTTGTTCCTTTAGTGAGGCTTAAT 980
Db 961 CTGCTCGAGACATCAAGTGGAGGGTAACTCTGCTAGCTCTGCAAGCTTCAAG 1020
QY 981 TAGATCCCATCGTCAATTTTACGACACTATCTTTCTAGGGTTAACTAGTCTGATCAG 1040
Db 1021 AAAATGCGGATCATCTCGCAAGAGAGATCTCTACTTTCTCCCTTTGCAAAACCAAGTTC 1080
QY 1041 GATCATATCTGCGGTCTTTTTCGCGCTCAGTATCGCCCAAGCTGCGCTATCTGGGC 1100
Db 1081 GACAACTCGTACGCGCTGTTGCAAGATCTACACCGCTCTGCAAGTGCCTCATCCA 1140
QY 1101 ATCGGGAGGAAGAGCGCGCTCTTTT-----CCCGCGAGGTTGAAGCGGCTGGAAGAGT 1158
Db 1141 AGGCGCAATCCTGATCCAAACCTTTTACTCCAGCGCCAGTAGGCGCTCTTTAAAGC 1200
QY 1159 TTGCGGAGATGACTGCTGCTGCAATGA----- 1186
Db 1201 TTGACCGAGAGCAATCCCGCAGTCTTCAGTGTGTGATGCTCTATGTTAGTCAACC 1260
QY 1187 ----- 1186
Db 1261 AATGCACTCAAGATTAGCGACCGAGCGGAATGCTTGGCCAGAGCATGTATCATATGGTTC 1320
QY 1187 GCTTACGCGGAAACCGACGCTTTTACCATGATTCGGGAAGGTGTGGCCATACGACGCTTT 1246
Db 1321 CAGAAACCTTATATCTGTTGGAAGTAACTACTTGCATGCTGTTGCTGCTGCTGCTGCTGCT 1380
QY 1247 TAAAGGTGAACTGTTGCTGTTAGGCGCACTGGGATACAGTTCGTCGCGGCTTTTCGGGAC 1306
Db 1381 TGCTTCTGCTCTTTTCTGGGAGATCGAGTGTCTATCGTAGGGGACCAACCTTTAA 1440
QY 1307 ACAGTTCGGATGCTGACCGGAGCGCATCAGCAACCGCAACATACCGGCGAGCGCG 1366
Db 1441 AGAGATCGCAATCTGAAATCTTGGTTTCAATTTGTAATACGTTTACTAGGCTTTCTGCTC 1500
QY 1367 GAACCTGCGCTGCGGTGTCAGATT----- 1391
Db 1501 TGTCACTTTGCTGCTGCTTATCTTTCCTGCTCATTTTTCGAGATAATTCGAGAAAT 1560
QY 1392 ----- 1391

Db 1561 CACATTACTTTATATAATATTAATTAATTCATTATGTGATAATCGGAATCGCTTAAGAAAAA 1620
Qy 1392 ----- 1391
Db 1621 AAGAGTCAATCCGTAGCGGAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1680
Qy 1392 ----- 1391
Db 1681 TAGAGTGTACTAGAGGAGGCCAAGAGTAATAGAAAAAAGAAAAATTGCGGAAAGAGCTGTG 1740
Qy 1392 ----- 1391
Db 1741 TTATGACTTCCCTGACTAATGCGGTGTTCAACGATACCTGGCAGTGACTCCTTAGCGCTC 1800
Qy 1392 -----AATGACAGCGGTGCGCGCTGGGATATTACGTACAGCGAGGACGGGTA 1438
Db 1801 ACCAAGCTCTTAAAAACGCGCGCTGACGAAATTCGATATCAAGCTTATCGATACCGTCG 1860
Qy 1439 TCCTGGCTGGATCGCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGCGGCGCGC 1498
Db 1861 ACCTGAGGGGGCGCGGTACCCAGCTTTTGTTCCTTTTGTGAGGGTTAAATTGCGCGC 1920
Qy 1499 TTGCGGTAAATCATGGTCAATAGCTGTTTCCTGTGTGAAATTTGTTATCCGCTCACAATTTCCA 1558
Db 1921 TTGCGGTAAATCATGGTCAATAGCTGTTTCCTGTGTGAAATTTGTTATCCGCTCACAATTTCCA 1980
Qy 1559 CACACATACGAGCGGGAACATAAAGTGTAAAGCCTTGGGGTGCTTAATGAGTGAGCTAA 1618
Db 1981 CACACATACGAGCGGGAACATAAAGTGTAAAGCCTTGGGGTGCTTAATGAGTGAGCTAA 2040
Qy 1619 CTCACATTAATTCGGTTGCGCTCACTGCGCGCTTTCAGTCGCGAAACCTGTCGCGCAG 1678
Db 2041 CTCACATTAATTCGGTTGCGCTCACTGCGCGCTTTCAGTCGCGAAACCTGTCGCGCAG 2100
Qy 1679 CTCGATTATGATCGGCCAACCGCGGGAGAGCGGTTTGGTATTTGGCGCTCTTCC 1738
Db 2101 CTGCAATTAATGAAATCGGCAACCGCGGGAGAGCGGTTTGGTATTTGGCGCTCTTCC 2160
Qy 1739 GCTTCTCGCTCACTGACTCGCTCGCTCGGTGCTTGGCTGCGGCGAGCGGTATCAGCT 1798
Db 2161 GCTTCTCGCTCACTGACTCGCTCGCTCGGTGCTTGGCTGCGGCGAGCGGTATCAGCT 2220
Qy 1799 CACTAAAGCGGTAAATACGGTTATCCAGAAATCAGGGGATAAACGAGAAAGAAACATG 1858
Db 2221 CACTAAAGCGGTAAATACGGTTATCCAGAAATCAGGGGATAAACGAGAAAGAAACATG 2280
Qy 1859 TGAGCAAAAGGCCAGCAAAAGGCCAGAACCTTAAAGCGCGGTTGCTGGCGTTTTC 1918
Db 2281 TGAGCAAAAGGCCAGCAAAAGGCCAGAACCTTAAAGCGCGGTTGCTGGCGTTTTC 2340
Qy 1919 CATAGGCTCGCGCCCTCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGA 1978
Db 2341 CATAGGCTCGCGCCCTCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGA 2400
Qy 1979 AACCCGACAGGACTATAAGATACAGGCGTTTCCCGCTGGAAAGTCCCTCGTGGCGCTC 2038
Db 2401 AACCCGACAGGACTATAAGATACAGGCGTTTCCCGCTGGAAAGTCCCTCGTGGCGCTC 2460
Qy 2039 CCGTTTCCGACCTCGCGCTTACCGGATACCTGTCGCGCTTTCCTCCCTTGGGAAGGCTG 2098
Db 2461 CCGTTTCCGACCTCGCGCTTACCGGATACCGGATACCGGATACCGGATACCGGATACCGG 2520
Qy 2099 GCGCTTCTCATAGCTCACGCTGTAGGTATCTCAGTTTCGGTGTAGGTGCTTCCCTCCAAG 2158
Db 2521 GCGCTTCTCATAGCTCACGCTGTAGGTATCTCAGTTTCGGTGTAGGTGCTTCCCTCCAAG 2580
Qy 2159 CTGGGCTGTGTGACGAACCCCGTTACGCCGACCGCTGCGCTTATCCCGTAACAT 2218
Db 2581 CTGGGCTGTGTGACGAACCCCGTTACGCCGACCGCTGCGCTTATCCCGTAACAT 2640
Qy 2219 CGTCTTGTAGTCCAAACCGGTAAAGACGACTTATCGCACTGCGCAGGCACTGGTAAC 2278

Db 2641 CGTCTTGAGTCCAAACCGGTAAAGACAGGACTTATCGCCTACTGCGCAGCAGCACTGGTAAC 2700
Qy 2279 AGGATTAGCAGAGCGAGGTATGTAGCGGTCTACAGAGTTCTTGAAGTGGTGGCTTAAC 2338
Db 2701 AGGATTAGCAGAGCGAGGTATGTAGCGGTCTACAGACTTCTTGAAGTGGTGGCTTAAC 2760
Qy 2339 TAGGGCTACACTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAGCCAGTACCTTC 2398
Db 2761 TAGGGCTACACTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAGCCAGTACCTTC 2820
Qy 2399 GGAAGAAAGAGTTGGTAGCTCTTGATCCGCAAAACCAACCCCTGCTAGCGGTGGTTTT 2458
Db 2821 GGAAGAAAGAGTTGGTAGCTCTTGATCCGCAAAAGAAACCAACCCCTGCTAGCGGTGGTTTT 2880
Qy 2459 TTTGTTTCAAGCAGCAGATTAACGCGCAGAAAAAGGATCTCAAGAAAGATCTTTTGATC 2518
Db 2881 TTTGTTTCAAGCAGCAGATTAACGCGCAGAAAAAGGATCTCAAGAAAGATCTTTTGATC 2940
Qy 2519 TTTTCTACGGGTCTGACGCTCAGTCGAAACGAAACTCAGTTTAAGGATTTTGGTCATG 2578
Db 2941 TTTTCTACGGGTCTGACGCTCAGTCGAAACGAAACTCAGTTTAAGGATTTTGGTCATG 3000
Qy 2579 AGATTATCAAAAGAGTCTTCACTAGATCTTTTAAATTTAAATTTAAATTTAAATTTAAATCA 2638
Db 3001 AGATTATCAAAAGAGTCTTCACTAGATCTTTTAAATTTAAATTTAAATTTAAATTTAAATCA 3060
Qy 2639 ATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTTACAAATGCTTAAATCAGTGAGGCA 2698
Db 3061 ATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTTACAAATGCTTAAATCAGTGAGGCA 3120
Qy 2699 CCTATCTCAGCGATCTGTCTATTTCTGTTTCATTCATAGTTGCTTCCCGCTCGTGTAG 2758
Db 3121 CCTATCTCAGCGATCTGTCTATTTCTGTTTCATTCATAGTTGCTTCCCGCTCGTGTAG 3180
Qy 2759 ATAACTACGATACGGGAGGCTTACCATCTGGCCAGTGTGCAATGATACCCGAGAC 2818
Db 3181 ATAACTACGATACGGGAGGCTTACCATCTGGCCAGTGTGCAATGATACCCGAGAC 3240
Qy 2819 CCACGCTCACCGCTCCAGATTTATCAGCAATAAACACGACGCGGAGGCGCGAGCGC 2878
Db 3241 CCACGCTCACCGCTCCAGATTTATCAGCAATAAACACGACGCGGAGGCGCGAGCGC 3300
Qy 2879 AGAAGTGTCTGCAACTTTATCCGCTCGATCCAGTCTTAAATTTGTTGCGCGGAGCT 2938
Db 3301 AGAAGTGTCTGCAACTTTATCCGCTCGATCCAGTCTTAAATTTGTTGCGCGGAGCT 3360
Qy 2939 AGAAGTGTCTGCGCAGTAAATAGTTTGCAGAGTTGTTGCCATTTGCTACAGGCATC 2998
Db 3361 AGAAGTGTCTGCGCAGTAAATAGTTTGCAGAGTTGTTGCCATTTGCTACAGGCATC 3420
Qy 2999 GTGGTGTACGCTCGTCTGTTGGTATGGCTTCAATTCAGCTCCGTTCCCAACGATCAAGG 3058
Db 3421 GTGGTGTACGCTCGTCTGTTGGTATGGCTTCAATTCAGCTCCGTTCCCAACGATCAAGG 3480
Qy 3059 CGAGTTACATGATCCCGCATTTGTCGCAAAAAGCGGTTAGCTTCCCTCGTCTCCGATC 3118
Db 3481 CGAGTTACATGATCCCGCATTTGTCGCAAAAAGCGGTTAGCTTCCCTCGTCTCCGATC 3540
Qy 3119 GTTGTGAGAAAGTAAAGTTGGCCGAGTGTATCACTCATGTTATGCGACACTGCAATAAT 3178
Db 3541 GTTGTGAGAAAGTAAAGTTGGCCGAGTGTATCACTCATGTTATGCGACACTGCAATAAT 3600
Qy 3179 TCTCTTACTGTATGCGCATCCGTAAGATCGTTTTCTGTAAGTGTGAGTACTCAACCAAG 3238
Db 3601 TCTCTTACTGTATGCGCATCCGTAAGATCGTTTTCTGTAAGTGTGAGTACTCAACCAAG 3660
Qy 3239 TCAATCTGAGAAATAGTGTATGCGGACCGAGTGTCTTTCGCGCGGCTCAATACGGGAT 3298
Db 3661 TCAATCTGAGAAATAGTGTATGCGGACCGAGTGTCTTTCGCGCGGCTCAATACGGGAT 3720
Qy 3299 AATACCGGCCCATAGCAGAACTTTAAAGTGTCTCATCTTGGAAAAAGCTTCTTCGGGG 3358
Db 3721 AATACCGGCCCATAGCAGAACTTTAAAGTGTCTCATCTTGGAAAAAGCTTCTTCGGGG 3780

QY 3359 CGAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCAGTGTAAACCCACTCGTGCA 3418
DB 3781 CGAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCAGTGTAAACCCACTCGTGCA 3840
QY 3419 CCGAACTGATCTTCAGCATCTTTTACTTTCACACAGGTTTCTGGGTGAGCAAAAACAGGA 3478
DB 3841 CCGAACTGATCTTCAGCATCTTTTACTTTCACACAGGTTTCTGGGTGAGCAAAAACAGGA 3900
QY 3479 AGGCAAAATGCCGCAAAAAGGGAATAAGGGCGACACGGAATGTTGAATATCTCATATC 3538
DB 3901 AGGCAAAATGCCGCAAAAAGGGAATAAGGGCGACACGGAATGTTGAATATCTCATATC 3960
QY 3539 TTCTCTTTTCAATATTAATTGAAGCATTTATCAGGGTTATGTTCTCATGAGCGGATACATA 3598
DB 3961 TTCTCTTTTCAATATTAATTGAAGCATTTATCAGGGTTATGTTCTCATGAGCGGATACATA 4020
QY 3599 TTGGAATGATTTAGAAAAATAAACAATAGAGGTTTCGCGCACATTTCCCGAAAAAGTG 3658
DB 4021 TTGGAATGATTTAGAAAAATAAACAATAGAGGTTTCGCGCACATTTCCCGAAAAAGTG 4080
QY 3659 CCAC 3662
DB 4081 CCAC 4084

RESULT 9
US-09-993-170-1
; Sequence 1, Application US/099931170
; Patent No. 6743969
; GENERAL INFORMATION:
; APPLICANT: Valent, Barbara
; APPLICANT: Bryant, Gregory
; TITLE OF INVENTION: Modification of a Plant Disease Resistance Gene
; TITLE OF INVENTION: Specificity and Method for Engineering Altered
; TITLE OF INVENTION: Specificity
; FILE REFERENCE: BB1385 US NA
; CURRENT APPLICATION NUMBER: US/09/993,170
; PRIOR FILING DATE: 2001-11-26
; PRIOR APPLICATION NUMBER: 60/248,335
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 4119
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Plaemid

US-09-993-170-1
Query Match 62.2%; Score 2279.2; DB 4; Length 4119;
Best Local Similarity 76.6%; Pred. No. 0;
Matches 3151; Conservative 0; Mismatches 508; Indels 457; Gaps 11;

QY 1 CTAATTGTAAGCGTTAAATTTTGTAAATTCGCGTTAAATTTTGTAAATTCAGTCT 60
DB 4 CTAATTGTAAGCGTTAAATTTTGTAAATTCGCGTTAAATTTTGTAAATTCAGTCT 63
QY 61 ATTTTAAACCAATAGGCGCAATCGGCAAAATCCCTTATAAATCAAAAGATAGACCGA 120
DB 64 ATTTTAAACCAATAGGCGCAATCGGCAAAATCCCTTATAAATCAAAAGATAGACCGA 123
QY 121 GATAGGTTGAGTGTTCAGTTTGGAAACAGAGTCCACTATTAAAGAGTGGAGTCT 180
DB 124 GATAGGTTGAGTGTTCAGTTTGGAAACAGAGTCCACTATTAAAGAGTGGAGTCT 183
QY 181 CAACGTCAAAAGGCGCAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCAC 240
DB 184 CAACGTCAAAAGGCGCAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCAC 243
QY 241 CTAATCAAGTTTTTGGGGTTCAGGTGCGTAAAGCACTAAATTCGGAACCCCTAAAGGGAG 300

DB 244 CTAATCAAGTTTTTGGGGTTCAGGTGCGGTAAAGCACTAAATTCGAAACCCCTAAAGGGAG 303
QY 301 CCCCCGATTTAGAGCTTCACGGGGAAAGCCGCGCAAACTGGCGAGAAAGGAAGGGAAGAA 360
DB 304 CCCCGATTTAGAGCTTCACGGGGAAAGCCGCGCAAACTGGCGAGAAAGGAAGGGAAGAA 363
QY 361 AGCGAAAGAGCGGCGCTAGGGCGCTGCAAGTGTAGCGGTACGCTGCGCGTAAACCAAC 420
DB 364 AGCGAAAGAGCGGCGCTAGGGCGCTGCAAGTGTAGCGGTACGCTGCGCGTAAACCAAC 423
QY 421 CACACCCCGCGCTTAATGCGCGCTACAGGGCGGCTCCCATTCGCCATTTCAGGCTGCG 480
DB 424 CACACCCCGCGCTTAATGCGCGCTACAGGGCGGCTCCCATTCGCCATTTCAGGCTGCG 483
QY 481 CAACGTGTTGGGAAGGCGCATCGGTGCGGCGCTCTTCGCTATTACGCCAGTGGCGAAAGG 540
DB 484 CAACGTGTTGGGAAGGCGCATCGGTGCGGCGCTCTTCGCTATTACGCCAGTGGCGAAAGG 543
QY 541 GGGATGTCGTCGAAGGCGATTAAAGTTGGGTAAAGCCAGGGTTTCCCAAGTCAACGCTTG 600
DB 544 GGGATGTCGTCGAAGGCGATTAAAGTTGGGTAAAGCCAGGGTTTCCCAAGTCAACGCTTG 603
QY 601 TAAACGAGCGCCAGTGA-----GCGCGCTCGTT 630
DB 604 TAAACGAGCGCCAGTGAATTGTAATACGACTCACTATAGGCGCAATTGGGTACCGGGCC 663
QY 631 CATTCAACGTTTTTGAACCCGTTGGAGGACGGGCGAGACTCGCGTGCATAATGCTTTTACAG 690
DB 664 CCCCTCGAGGTGCGATCCGAGGTAGGTCTAGGGGCGCTGATCCTCACAATATTTTGTAA 723
QY 691 CGTGATGGAGCAGATGAAGATGCTCGACACGCTCGAGAACACGCGAGCT----- 738
DB 724 ATTTCAAAAGTCAGGAGCATGAATTATGTAAGTTAATAATATGGGCCCAACTCTTAC 783
QY 739 -----AGATTAACCTAGAAAGATAATCATATGTCAGTGTAAAGATATCAT 790
DB 784 CTTATATAAAATTTGGGATGATATATAATAAAGTGGACCTTAATTA CTTGATATAAT 843
QY 791 GCGTAAATTCAGCATGGGATCTGTAATACGACTCACTA--TAGGGCGAATGGGTACC 848
DB 844 GCAGATATTAACACTAGCAAAATATAATTCGATAATATTAATGCTTAATACGCAT 903
QY 849 GGGCCCCCTCGAGGTGCGAGGTATCGATAAGCTTGATCGAATTCCTGAGCGCCGGG 908
DB 904 TAATAAACCAATAAGTTTTACATCTTCTAAAGCTTTGAAAAAAGTCAAGCTGAATAA 963
QY 909 GGATCCACTAGTTCTAGAGCGGCGCCACCGGGTGGAGCTCCAGCTTTTGTCCCTTTA 968
DB 964 TAAATAAGTTGGCGTTGTTATAAATCGACCGGTTTCGCGCTTTATGGTTTAAATTCGA 1023
QY 969 GTGAG-----GGTTAATTAGATCCCATCGCTCAATTTTACGCAGACTATCTTTCT 1018
DB 1024 TAGAACAATTTTGTGTTAATTCGAACATACAACGAATATTCACACTGACTGAAAAATCG 1083
QY 1019 AGGGTAAATCTAGTGTGATCAGGATCATA----- 1047
DB 1084 ACAGTTTTTGTGCAATCAACATTAATAATTAATAAATAAATAAATAAATAAATAA 1143
QY 1048 -----TCGTCGCGCTTTTTCGCGCTCAGTCATCGCCCAAGCTGGCGCTATCTGG 1099
DB 1144 TTATTTTGTGATCGATGCTTTTTTATTCATTTTTCACACCGGTTGCGATTTCGG 1203
QY 1100 CAT----- 1102
DB 1204 CTTCAACCAATTTGGGACCTTTTTCACACCGGTTTACGATTACAAATCCAATTCCAAACC 1263
QY 1103 -----CGGGAGGAAGAACCCGCTGCTTT----- 1127
DB 1264 ATATCCAGGAGATTTAAAAAGGCGGCTTATTAATGAACGCTTATTCCTCAATGTTTCA 1323
QY 1128 -----TCCCGGAGGTTGAAGGCGATGGAAGAGTTTGGCGGAGTACGCTGCTGCTCA 1182
DB 1324 CGAGGCTTCCGAATTCGTGCGCGCTTAAAAAGTTGTGCGGAGCTCGCTCGTGGGCT 1383

Qy	1183	TTGAGCTGTGAGCGAAACGGCA--CGTTTACCATGATGATTCGGGAAG-----	1227
Db	1384	ATCAGCGCGTTTAAAAAATGCACAATCGGTATTATTAGATTAAATCTTTAAAACTGACGACAG	1443
Qy	1228	-----	1227
Db	1444	ATATTCAAAACTGGGTTCAAAAGAAATTTTAAACGAAATTTACAGGAATGTAACAGGAGCG	1503
Qy	1228	-----GTGTGGCCATGACGCGCTTTTAAACGGTGAACCTGTTCGTTTCAGGCCACCT	1275
Db	1504	CGGACGAAATTTCTCTAACCTGCGCACGATAAAATGTTTATACGTGCGTCCGAGAAGGAG	1563
Qy	1276	GGGATACAGTTTCGTCCGGCGCTTTTCGGGACACAGTTCCGGATCGGTCAAGCCCGAAGCGCA	1335
Db	1564	TTTCATAATTTGGCGTATGCACTTATTAAACGAAAGAATAATGTTTATATGCCCTCTTTCT	1623
Qy	1336	TCAGCAACCCGGAACAATACCGCGGACAC-----	1362
Db	1624	TCAAACCCCGGTAACACAGAGGGAATTAATGCGCGGTAAACCAAGATACAGTTATTATTAC	1683
Qy	1363	-----GCCGGAACTGCCGTGCCGGTG	1383
Db	1684	ATGAATGGTGCATATAATTTTAAACAGATGGAAGATTATGGTTACGAATGGATGGGA	1743
Qy	1384	TGCAGATTAAATGACAGCGGTGCGCGCTGGGATATTAGCTCAGCGGAGACGGGTATCTCTG	1443
Db	1744	TTCACAAAATTTGGATAGTACAGAAAGTATTAAAAACCCCGACAGTTATGCTATTTTGTGCAC	1803
Qy	1444	GCTGGAATGCCGACAGAAATGGACATGGATACCCCGTGAGTTTACCCGGCGGG-----	1493
Db	1804	AAATGTCACGTTTATAAATATTGTTAAATACGTTAGTGTGGAAATGAGAGGAATCGCGGT	1863
Qy	1494	-----	1493
Db	1864	CAATCTAACAATAGAGGGATCCACTAGTCTTAGAGCGCGCCACCGCGTGGAGCTCCA	1923
Qy	1494	-----CGCGCTTGGCGTAAATCATGCTCATAGCTGT	1523
Db	1924	GCTTTTGTTCCTTTTAGTAGGGTTAATTTGAGAGCTTTGGCGTAAATCATGCTCATAGCTGT	1983
Qy	1524	TTCTGTGTGAAATTTGTTATTCGCTCACAATTTCCACACAAATACGAGCGCGGAAGCATAA	1583
Db	1984	TTCTGTGTGAAATTTGTTATTCGCTCACAATTTCCACACAAATACGAGCGCGGAAGCATAA	2043
Qy	1584	AGTGTAAAGCTGGGGTGCCCTAATGAGTAGCTAACTCAATTAATTTGCGTTGCGCTCAC	1643
Db	2044	AGTGTAAAGCTGGGGTGCCCTAATGAGTAGCTAACTCAATTAATTTGCGTTGCGCTCAC	2103
Qy	1644	TGCCCCGTTTCCAGTCGGGAAAACCTGTTCGTGCCAGCTGCATTAATGAATCGGCCCAAGCG	1703
Db	2104	TGCCCCGTTTCCAGTCGGGAAAACCTGTTCGTGCCAGCTGCATTAATGAATCGGCCCAAGCG	2163
Qy	1704	CGGGAGAGCGGTTTCGCTATTGGGCGCTCTTCGCGCTTCCTCGCTCACTGACTCGCTGC	1763
Db	2164	CGGGAGAGCGGTTTCGCTATTGGGCGCTCTTCGCGCTTCCTCGCTCACTGACTCGCTGC	2223
Qy	1764	GCTCGTGTGTTTCGCTCGCGGACGCGTATCAGCTCACTCAAAAGCGGGTAATACGGTTAT	1823
Db	2224	GCTCGTGTGTTTCGCTCGCGGACGCGTATCAGCTCACTCAAAAGCGGGTAATACGGTTAT	2283
Qy	1824	CCACAGAAATCAGGGGATAACGCGAGGAAGACATGTCAGCAAAAGGCCACCAAAAGGCCA	1883
Db	2284	CCACAGAAATCAGGGGATAACGCGAGGAAGACATGTCAGCAAAAGGCCACCAAAAGGCCA	2343
Qy	1884	GGAAACCGTAAAAAGCCCGCTGTGCTGCGGTTTTTTCATATAGGCTCCGCCCCCTCGACGAGC	1943
Db	2344	GGAAACCGTAAAAAGCCCGCTGTGCTGCGGTTTTTTCATATAGGCTCCGCCCCCTCGACGAGC	2403
Qy	1944	ATCACAAAAATTCGACGCTCAAGTCAGAGGTGGGAAAACCCGACAGGACTATAAGATACC	2003
Db	2404	ATCACAAAAATTCGACGCTCAAGTCAGAGGTGGGAAAACCCGACAGGACTATAAGATACC	2463

QY	2004	AGGCGTTTCCCGCTGGGAAGCTCCGTCGTGGCGCTCTCCTGTTCCGACCGCTGCCGCTTACCG	2063
DB	2464	AGGCGTTTCCCGCTGGGAAGCTCCGTCGTGGCGCTCTCCTGTTCCGACCGCTGCCGCTTACCG	2523
QY	2064	GATACCTGTCCGCGCTTTCCTCCCTTTCGGGAAGCGTGGCGCTTCTCATAGCTCACGCTGTA	2123
DB	2524	GATACCTGTCCGCGCTTTCCTCCCTTTCGGGAAGCGTGGCGCTTCTCATAGCTCACGCTGTA	2583
QY	2124	GGTATCTCAGTTCCGGTGTAGGTGGTTCCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCG	2183
DB	2584	GGTATCTCAGTTCCGGTGTAGGTGGTTCCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCG	2643
QY	2184	TTGAGCCCGAGCGCTGGCGCTTATTCCGGTAACTATCGTCTTGAGTCCAACCCCGGTAAAGAC	2243
DB	2644	TTGAGCCCGAGCGCTGGCGCTTATTCCGGTAACTATCGTCTTGAGTCCAACCCCGGTAAAGAC	2703
QY	2244	ACGACTTATCGCCACTGGCAGCAGCCACTCGTAAACAGGATTACGACGAGGAGTATGTAG	2303
DB	2704	ACGACTTATCGCCACTGGCAGCAGCCACTCGTAAACAGGATTACGACGAGGAGTATGTAG	2763
QY	2304	CGGCTGCTACAGAGTTCTTCAAGTGGTGGCTTAACCTACGGCTACACTAGAAGGACAGTAT	2363
DB	2764	CGGCTGCTACAGAGTTCTTCAAGTGGTGGCTTAACCTACGGCTACACTAGAAGGACAGTAT	2823
QY	2364	TTGGTATCTCGGCTCTGCTGAAACGACTTACCTTCGGAAGGAGTGGTAGCTCTTGAT	2423
DB	2824	TTGGTATCTCGGCTCTGCTGAAACGACTTACCTTCGGAAGGAGTGGTAGCTCTTGAT	2883
QY	2424	CCGCAACAAACCCACCGCTGGTAGCGGTGGTTTTTTTGTTCGACGACGAGATTACCG	2483
DB	2884	CCGCAACAAACCCACCGCTGGTAGCGGTGGTTTTTTTGTTCGACGACGAGATTACCG	2943
QY	2484	GCAGAAAAAAGGATCTCAAGAAAGATCCTTTTGATCTTTTCTACGGGCTCTGACGCTCAGT	2543
DB	2944	GCAGAAAAAAGGATCTCAAGAAAGATCCTTTTGATCTTTTCTACGGGCTCTGACGCTCAGT	3003
QY	2544	GGAAACGAAAACTCA CGTTAAAGGATTTGGTCAAGATTTATCAAAAGGATCTTCACCT	2603
DB	3004	GGAAACGAAAACTCA CGTTAAAGGATTTGGTCAAGATTTATCAAAAGGATCTTCACCT	3063
QY	2604	AGATCCTTTTAAATTAATAATGAAGTTTAAATCAATCTAAAGTATATATGAGTAACTT	2663
DB	3064	AGATCCTTTTAAATTAATAATGAAGTTTAAATCAATCTAAAGTATATATGAGTAACTT	3123
QY	2664	GGTCTGACAGTTACCAATGCTTAATCAGTCAGGACCTATCTCAGCGATCTGCTATTTTC	2723
DB	3124	GGTCTGACAGTTACCAATGCTTAATCAGTCAGGACCTATCTCAGCGATCTGCTATTTTC	3183
QY	2724	GTTCCATCCATAGTTGCTGACTCCCGTCGTGTAGATAACTACGATACGGAGGGCTTAC	2783
DB	3184	GTTCCATCCATAGTTGCTGACTCCCGTCGTGTAGATAACTACGATACGGAGGGCTTAC	3243
QY	2784	CATCTGCGCCAGTGTGCAATGATACCGCGAGACCA CGCTCACCGGCTCCAGATTTAT	2843
DB	3244	CATCTGCGCCAGTGTGCAATGATACCGCGAGACCA CGCTCACCGGCTCCAGATTTAT	3303
QY	2844	CAGCAATAAACACGACCGCGAAGCGCGAGCGCAGAAAGTGTCTCTGCAACTTTATCCG	2903
DB	3304	CAGCAATAAACACGACCGCGAAGCGCGAGCGCAGAAAGTGTCTCTGCAACTTTATCCG	3363
QY	2904	CCTCCATCCAGTCTATTAAATTGTTGGCGGAAGCTAGAGTAAGTAGTTCGCCAGTTAAATA	2963
DB	3364	CCTCCATCCAGTCTATTAAATTGTTGGCGGAAGCTAGAGTAAGTAGTTCGCCAGTTAAATA	3423
QY	2964	GTTTGGCGAACGTTGTGCAATTGCTACAGGCAATCGTGGTGTCA CGCTCGTCTGTTGGTA	3023
DB	3424	GTTTGGCGAACGTTGTGCAATTGCTACAGGCAATCGTGGTGTCA CGCTCGTCTGTTGGTA	3483
QY	3024	TGGCTTCAATTCAGTCCGTTCCCAACGATCAAGCGAGGATTAATCATCATCCCCCATCTGT	3083
DB	3484	TGGCTTCAATTCAGTCCGTTCCCAACGATCAAGCGAGGATTAATCATCATCCCCCATCTGT	3543
QY	3084	GCAAAAAGCGGTTAGTCTCTTCGGTCTCCGATCGTTGTTCAGAAAGTTCGCGCCAG	3143

Db	3544	GCATAAAAGCGGTAGCTCCTTCGGTCTCCGATCGTTCAGAAAGTAAGTTGGCGCAG	3603	
Qy	3144	TGTTATCACTCATGGTTATGGCAGCACTGCATAATCTCTTACTGTCAATGCCATCGGTAA	3203	
Db	3604	TGTTATCACTCATGGTTATGGCAGCACTGCATAATCTCTTACTGTCAATGCCATCGGTAA	3663	
Qy	3204	GATGCTTTTCTGTGACTGGTGAAGTACTCAACCAAGTCAATCTGAGAAATAGTGTATGCGGC	3263	
Db	3664	GATGCTTTTCTGTGACTGGTGAAGTACTCAACCAAGTCAATCTGAGAAATAGTGTATGCGGC	3723	
Qy	3264	GACCGAGTTGCTCTTGGCCCGGCTCAATACGGGATAATACCGGGCCACATAGCAAACTT	3323	
Db	3724	GACCGAGTTGCTCTTGGCCCGGCTCAATACGGGATAATACCGGGCCACATAGCAAACTT	3783	
Qy	3324	TAAAGTGTCTCATCTGGAAGAACGTTCTTCGGGGGGAACCTCTCAAGGATCTTACCGC	3383	
Db	3784	TAAAGTGTCTCATCTGGAAGAACGTTCTTCGGGGGGAACCTCTCAAGGATCTTACCGC	3843	
Qy	3384	TGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCCAACTGATCTTCAGCATCTTTTA	3443	
Db	3844	TGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCCAACTGATCTTCAGCATCTTTTA	3903	
Qy	3444	CTTTCAACAGCTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCGCAAAAAGGGAA	3503	
Db	3904	CTTTCAACAGCTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCGCAAAAAGGGAA	3963	
Qy	3504	TAAGGGCAGACGGAAATGTTCAATACTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	3563	
Db	3964	TAAGGGCAGACGGAAATGTTCAATACTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	4023	
Qy	3564	TTTATCAGGGTTATGTTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAC	3623	
Db	4024	TTTATCAGGGTTATGTTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAC	4083	
Qy	3624	AAATAGGGTTCGGCGCACATTTCCCGGAAAGTGC	3659	
Db	4084	AAATAGGGTTCGGCGCACATTTCCCGGAAAGTGC	4119	
RESULT 10				
US-09-138-024-20				
; Sequence 20, Application US/09138024A				
; Patent No. 6004779				
; GENERAL INFORMATION:				
; APPLICANT: Bradley, John D.				
; APPLICANT: Thompson, Craig M.				
; APPLICANT: Moore, Jeffrey B.				
; APPLICANT: Wobbe, C. Richard				
; APPLICANT: Healy, Judith M.				
; APPLICANT: Donnelly, Caroline E.				
; TITLE OF INVENTION: REGULATED GENE EXPRESSION IN YEAST				
; FILE REFERENCE: 0342/1046SUS1				
; CURRENT APPLICATION NUMBER: US/09/138,024A				
; CURRENT FILING DATE: 1998-08-21				
; EARLIER APPLICATION NUMBER: 60/056,719				
; EARLIER FILING DATE: 1997-08-22				
; NUMBER OF SEQ ID NOS: 24				
; SOFTWARE: FastSeq for Windows Version 3.0				
; SEQ ID NO 20				
; LENGTH: 7102				
; TYPE: DNA				
; ORGANISM: Artificial Sequence				
; FEATURE:				
; OTHER INFORMATION: Plasmid p2M195				
US-09-138-024-20				
Query Match 59.3%; Score 2172.8; DB 3; Length 7102;				
Best Local Similarity 99.7%; Pred. No. 0;				
Matches 217; Conservative 7; Mismatches 0; Indels 0; Gaps 0;				
Qy	1479	GAGTTACCGCGGGCGCGCTTGGCGTAATCATGTCATAGCTGTTCTCTGTGAAATTT	1538	

Db	3341	GAGTCGACGGGCATGCAAGCTTTGGCGTAATCATGGTCATAGCTGTTTCTGTGTGAAATTT	3400	
Qy	1539	GTTATCCGCTCACAATTCACACAACATACGAGCGGAGCATAAAGTGTAAAGCCTGGG	1598	
Db	3401	GTTATCCGCTCACAATTCACACAACATACGAGCGGAGCATAAAGTGTAAAGCCTGGG	3460	
Qy	1599	GTGCTTAATGAGTGAGCTAACTCACATTAATTTGCGTTGCGCTCACTGCCCCGTTTCCAGT	1658	
Db	3461	GTGCTTAATGAGTGAGCTAACTCACATTAATTTGCGTTGCGCTCACTGCCCCGTTTCCAGT	3520	
Qy	1659	CGGGAAAACCTGTGTGCGCAGCTGCATTAATGAATCGGCGCAACGCGCGGGGAGAGCGGTT	1718	
Db	3521	CGGGAAAACCTGTGTGCGCAGCTGCATTAATGAATCGGCGCAACGCGCGGGGAGAGCGGTT	3580	
Qy	1719	TGCGTATTTGGGCGCTCTTCGCTTCCGCTTCCGCTCACTGCTCACTGCTCGGTCGTTCCGC	1778	
Db	3581	TGCGTATTTGGGCGCTCTTCGCTTCCGCTTCCGCTCACTGCTCACTGCTCGGTCGTTCCGC	3640	
Qy	1779	TGCGCGAGCGGTATCAGTCACTCAAAAGGCGGTAAATACGGTTATCCACAGAAATCAGGG	1838	
Db	3641	TGCGCGAGCGGTATCAGTCACTCAAAAGGCGGTAAATACGGTTATCCACAGAAATCAGGG	3700	
Qy	1839	ATAACGCGAGGAAAGAAATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCCGTAAAGG	1898	
Db	3701	ATAACGCGAGGAAAGAAATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCCGTAAAGG	3760	
Qy	1899	CGCGTTCGTCGCGCTTTTTCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGAC	1958	
Db	3761	CGCGTTCGTCGCGCTTTTTCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGAC	3820	
Qy	1959	GCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAGATACACAGCGTTTCCCGCTG	2018	
Db	3821	GCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAGATACACAGCGTTTCCCGCTG	3880	
Qy	2019	GAAGTCCCTGTGCGGCTCTCTGTTCCGACCCCTGCGCTTACCGGATACCTGTCGCGCT	2078	
Db	3881	GAAGTCCCTGTGCGGCTCTCTGTTCCGACCCCTGCGCTTACCGGATACCTGTCGCGCT	3940	
Qy	2079	TTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGG	2138	
Db	3941	TTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGG	4000	
Qy	2139	TGTAGTCTGTGCTCCAAAGCTGGGCTGTGTGCAGAAACCCCGCTTACGCGCGACCGCT	2198	
Db	4001	TGTAGTCTGTGCTCCAAAGCTGGGCTGTGTGCAGAAACCCCGCTTACGCGCGACCGCT	4060	
Qy	2199	GGCCTTATCCGTTAACHTATGTTGAGTCCAAACCCGCTGAGACACGACTTATCGCCAC	2258	
Db	4061	GGCCTTATCCGTTAACHTATGTTGAGTCCAAACCCGCTGAGACACGACTTATCGCCAC	4120	
Qy	2259	TGGCAGCAGCCACTGCTAACAGGATTTAGCAGAGCGAGGTATGTAGGCGGTGCTCAGAGT	2318	
Db	4121	TGGCAGCAGCCACTGCTAACAGGATTTAGCAGAGCGAGGTATGTAGGCGGTGCTCAGAGT	4180	
Qy	2319	TCTTGAAGTGTGGCTTAACCTACGCTACATAGAAAGACAGTATTGTTGATCTCGGCTC	2378	
Db	4181	TCTTGAAGTGTGGCTTAACCTACGCTACATAGAAAGACAGTATTGTTGATCTCGGCTC	4240	
Qy	2379	TGCTGAAGCGAGTTACCTTCGGNAAGAGTTGGTAGCTCTTGATCCGCGCAACCAACCA	2438	
Db	4241	TGCTGAAGCGAGTTACCTTCGGNAAGAGTTGGTAGCTCTTGATCCGCGCAACCAACCA	4300	
Qy	2439	CGCTGTGTAGGCTGGTGTGTTTTTTTGTTCGAAGCAGCAGATTTACGCGCAGAAAAAGGAT	2498	
Db	4301	CGCTGTGTAGGCTGGTGTGTTTTTTTGTTCGAAGCAGCAGATTTACGCGCAGAAAAAGGAT	4360	
Qy	2499	CTCAAGAAAGATCTTTTGATCTTTTCTACGGGTCGACGCTCAGTGGGAACGAAACTCAC	2558	
Db	4361	CTCAAGAAAGATCTTTTGATCTTTTCTACGGGTCGACGCTCAGTGGGAACGAAACTCAC	4420	
Qy	2559	GTTAAGGATTTGGTCAATGAGATTTATCAAAAGGATCTTCACTAGATCTCTTTTAAATTT	2618	
Db	4421	GTTAAGGATTTGGTCAATGAGATTTATCAAAAGGATCTTCACTAGATCTCTTTTAAATTT	4480	

QY 2079 TTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTGG 2138
Db 3941 TTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTGG 4000
QY 2139 TGTAGGTGTTCGCTCCAAAGCTGGGTGTGTGCACGAACCCCGCTTGACGCCGACCGCT 2198
Db 4001 TGTAGGTGTTCGCTCCAAAGCTGGGTGTGTGCACGAACCCCGCTTGACGCCGACCGCT 4060
QY 2199 GGGCTTATCCGGTAACATATCGTCTTGAGTCCAAACCCGGTAAGACACGACTTATCGCCAC 2258
Db 4061 GGGCTTATCCGGTAACATATCGTCTTGAGTCCAAACCCGGTAAGACACGACTTATCGCCAC 4120
QY 2259 TGGCAGCAGCCACTGGTAAACAGGATTAGCAGAGCGAGGTATGTAGCGGTGTCTACAGAGT 2318
Db 4121 TGGCAGCAGCCACTGGTAAACAGGATTAGCAGAGCGAGGTATGTAGCGGTGTCTACAGAGT 4180
QY 2319 TCTTGAAGTGTGGCTTAACCTAACGGGTACATAGAGGACAGTATTTGGTATCTCGGCTC 2378
Db 4181 TCTTGAAGTGTGGCTTAACCTAACGGGTACATAGAGGACAGTATTTGGTATCTCGGCTC 4240
QY 2379 TGCTGAAGCAGTACCTTCGGGAAAAGAGTTGGTGTAGCTCTTGATCCGGCAACCAACCA 2438
Db 4241 TGCTGAAGCAGTACCTTCGGGAAAAGAGTTGGTGTAGCTCTTGATCCGGCAACCAACCA 4300
QY 2439 CCGCTGGTAGCGGTGGTGTGTTTGTGCAAGCAGCAGATTACGGCGCAGAAAAAGGAT 2498
Db 4301 CCGCTGGTAGCGGTGGTGTGTTTGTGCAAGCAGCAGATTACGGCGCAGAAAAAGGAT 4360
QY 2499 CTCAGAAGATCCTTTGTGATCTTTTCTACGGGTCTGACGCTCAGTGGAAACGAAACTCAC 2558
Db 4361 CTCAGAAGATCCTTTGTGATCTTTTCTACGGGTCTGACGCTCAGTGGAAACGAAACTCAC 4420
QY 2559 GTTAAGGATTTTGGTCTAGATGATTATCAAAAGCATCTTCACTAGATCTCTTTAAAT 2618
Db 4421 GTTAAGGATTTTGGTCTAGATGATTATCAAAAGCATCTTCACTAGATCTCTTTAAAT 4480
QY 2619 AAAATGAAGTTTAAATCAATCTAAAGTATATATGAGTAAATCTGGTCTGACAGTTACC 2678
Db 4481 AAAATGAAGTTTAAATCAATCTAAAGTATATATGAGTAAATCTGGTCTGACAGTTACC 4540
QY 2679 AATGCTTAATCAGTGAGGACCTATCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTTG 2738
Db 4541 AATGCTTAATCAGTGAGGACCTATCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTTG 4600
QY 2739 CCTGACTCCCGTCTGTGTAGATAACTACGATACGGAGGGCTTACCATCTGSCCCAGTG 2798
Db 4601 CCTGACTCCCGTCTGTGTAGATAACTACGATACGGAGGGCTTACCATCTGSCCCAGTG 4660
QY 2799 CTGCAATGATACCGGAGACCCACGCTCACCGGCTCCAGATTATCAGCAATAAACCCAGC 2858
Db 4661 CTGCAATGATACCGGAGACCCACGCTCACCGGCTCCAGATTATCAGCAATAAACCCAGC 4720
QY 2859 CAGCCGAAGGCCGAGCGCAGAAAGTGTCTGTGCAACTTTTATCCGCTCCATCCAGTCTA 2918
Db 4721 CAGCCGAAGGCCGAGCGCAGAAAGTGTCTGTGCAACTTTTATCCGCTCCATCCAGTCTA 4780
QY 2919 TTAATTTGTCGGGAAGCTAGATGAGTGTTCGCGAGTTAATAGTTTGGCAAGTTG 2978
Db 4781 TTAATTTGTCGGGAAGCTAGATGAGTGTTCGCGAGTTAATAGTTTGGCAAGTTG 4840
QY 2979 TTGCCATTGCTACGGCATCGTGGTGTCACTCGCTCGTTTGGTATGGCTTCATTTCAGCT 3038
Db 4841 TTGCCATTGCTACGGCATCGTGGTGTCACTCGCTCGTTTGGTATGGCTTCATTTCAGCT 4900
QY 3039 CCGGTTCCCAACGATCAAGCGGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGTTA 3098
Db 4901 CCGGTTCCCAACGATCAAGCGGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGTTA 4960
QY 3099 GCTCCCTCGGCTCCGATCGTTGTGAGAGTGTGGCGCAGTGTATACATCATTG 3158
Db 4961 GCTCCCTCGGCTCCGATCGTTGTGAGAGTGTGGCGCAGTGTATACATCATTG 5020

QY 3159 TTATGGCAGCACTGCAATAATTCTCTTACTGTCAATGCCATCCGTAAGATGCTTTTCTGTGA 3218
Db 5021 TTATGGCAGCACTGCAATAATTCTCTTACTGTCAATGCCATCCGTAAGATGCTTTTCTGTGA 5080
QY 3219 CTGGTGTAGTACTCAACCAAGTCATCTGAGATAGTGTATGGCGGACCGAGTTGCTCTT 3278
Db 5081 CTGGTGTAGTACTCAACCAAGTCATCTGAGATAGTGTATGGCGGACCGAGTTGCTCTT 5140
QY 3279 GCCCGCTCAATACGGGATAATACCGGCCACATAGCAGAACTTTAAAAAGTGTCTCATCA 3338
Db 5141 GCCCGCTCAATACGGGATAATACCGGCCACATAGCAGAACTTTAAAAAGTGTCTCATCA 5200
QY 3339 TTGAAAAAGCTTCTTCGGGGGAAAACTCTCAAGATCTTACCGCTGTGAGATCCAGTT 3398
Db 5201 TTGAAAAAGCTTCTTCGGGGGAAAACTCTCAAGATCTTACCGCTGTGAGATCCAGTT 5260
QY 3399 CGATGTAAACCACTCGTGCACCCACCTGATCTTTCAGCATCTTTTACTTTTACCGGTTT 3458
Db 5261 CGATGTAAACCACTCGTGCACCCACCTGATCTTTCAGCATCTTTTACTTTTACCGGTTT 5320
QY 3459 CTGGTGTAGCAAAAAACAGGAAGGCAAAATGCGCAAAAAAGGAATAAGGGCGACACGGA 3518
Db 5321 CTGGTGTAGCAAAAAACAGGAAGGCAAAATGCGCAAAAAAGGAATAAGGGCGACACGGA 5380
QY 3519 AATGTGTAATCTCATCTCTTCTTTTCAATATTTGAAGCATTTATCAGGTTTATT 3578
Db 5381 AATGTGTAATCTCATCTCTTCTTTTCAATATTTGAAGCATTTATCAGGTTTATT 5440
QY 3579 GTCTCATGAGCGGATACATATTTGATCTATTTAGAAAATAAACAATAGGGTTCCGC 3638
Db 5441 GTCTCATGAGCGGATACATATTTGATCTATTTGAATGTATTTAGAAAATAAACAATAGGGTTCCGC 5500
QY 3639 GCACATTTCCCGAAAAAGTGCCAC 3662
Db 5501 GCACATTTCCCGAAAAAGTGCCAC 5524

RESULT 12
US-09-573-322-20
; Sequence 20, Application US/09573322
; Patent No.: 6531289
; GENERAL INFORMATION:
; APPLICANT: Bradley, John D.
; APPLICANT: Thompson, Craig M.
; APPLICANT: Moore, Jeffrey B.
; APPLICANT: Wobbe, C. Richard
; APPLICANT: Bailey, David A.
; TITLE OF INVENTION: Regulated Gene Expression in Yeast and
; FILE REFERENCE: 0342/1D469-US4
; CURRENT APPLICATION NUMBER: US/09/573,322
; PRIOR FILING DATE: 2000-05-18
; PRIOR APPLICATION NUMBER: 09/404,066
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 09/138,024
; PRIOR FILING DATE: 1998-08-21
; PRIOR APPLICATION NUMBER: 60/056,719
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 20
; LENGTH: 7102
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: ZM195 plasmid
US-09-573-322-20

Query Match 59.3%; Score 2172.8; DB 4; Length 7102;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2177; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 1479 GAGTTACCCGGGGCGCGCTTGGCGTAATCATGTGCTAGTGTCTCTGTGTAATT 1538

RESULT 13
US-09-138-024-21
; Sequence 21, Application US/09138024A
; Patent No. 6004779
; GENERAL INFORMATION:
; APPLICANT: Bradley, John D.
; APPLICANT: Thompson, Craig M.
; APPLICANT: Moore, Jeffrey B.
; APPLICANT: Wobbe, C. Richard
; APPLICANT: Healy, Judith M.
; APPLICANT: Donnelly, Caroline E.
; TITLE OF INVENTION: REGULATED GENE EXPRESSION IN YEAST
; FILE REFERENCE: 0342/1D469US1
; CURRENT APPLICATION NUMBER: US/09/138,024A
; EARLIER FILING DATE: 1998-08-21
; EARLIER APPLICATION NUMBER: 60/056,719
; EARLIER FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 21
; LENGTH: 7333
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Plasmid pZM197
US-09-138-024-21

		Query Match	59.3%;	Score 2172.8;	DB 3;	Length 7333;
		Best Local Similarity	99.7%;	Pred. No. 0;		
		Matches 2177;	Conservative 0;	Mismatches 7;	Indels 0;	Gaps 0;
QY	1479	GAGTTACCGCGCGCGCTGGCGTAATCATGCTCATAGCTGTTCCTGTGTGCAAT	1538			
DB	3576	GAGTCGACCGGATCAAGCTTGGCGTAATCATGCTCATAGCTGTTCCTGTGTGCAAT	3635			
QY	1539	GTTATCCGCTCACAAATTCACACACATACGAGCGGAGCATAAAGTGTAAGCCTGGG	1598			
DB	3636	GTTATCCGCTCACAAATTCACACACATACGAGCGGAGCATAAAGTGTAAGCCTGGG	3695			
QY	1599	GTGCTAATAGTAGCTAACTCAATTAATGCTGCTGCTCACTGCGGCTTTCAGT	1658			
DB	3696	GTGCTAATAGTAGCTAACTCAATTAATGCTGCTGCTCACTGCGGCTTTCAGT	3755			
QY	1659	CGGGAACCTGCTGCGGAGCTGCAATTAATGAAATCGGCGCAAGCGCGGAGAGCGGT	1718			
DB	3756	CGGGAACCTGCTGCGGAGCTGCAATTAATGAAATCGGCGCAAGCGCGGAGAGCGGT	3815			
QY	1719	TGCGTATTGGGCGCTCTTCGCTTCTCGCTCACTGACTCGCTGGCTCGGTTCGGC	1778			
DB	3816	TGCGTATTGGGCGCTCTTCGCTTCTCGCTCACTGACTCGCTGGCTCGGTTCGGC	3875			
QY	1779	TGCGCGAGCGGTATCAGCTCACTCAAAAGCGGTAATACGGTTATCCACAGATCAGGG	1838			
DB	3876	TGCGCGAGCGGTATCAGCTCACTCAAAAGCGGTAATACGGTTATCCACAGATCAGGG	3935			
QY	1839	ATAACGAGGAGAGATGTGAGCAAAAGGCGCAAAAGGCGGAGACCGTAAAGG	1898			
DB	3936	ATAACGAGGAGAGATGTGAGCAAAAGGCGCAAAAGGCGGAGACCGTAAAGG	3995			
QY	1899	CGCGCTGCTGGCGTTTTCATAGGCTCCGCCCCCTTGACGAGCATCACAAAATCGAC	1958			
DB	3996	CGCGCTGCTGGCGTTTTCATAGGCTCCGCCCCCTTGACGAGCATCACAAAATCGAC	4055			
QY	1959	GCTAAGTCAGAGTGGCGAAACCCGACAGGACTATAAGATACAGAGGTTTCCCTG	2018			
DB	4056	GCTAAGTCAGAGTGGCGAAACCCGACAGGACTATAAGATACAGAGGTTTCCCTG	4115			
QY	2019	GAAGTCCCTGTCGCTCTCTGTTCCGACCGCTGCGCTACCGGATACCTGTCCGCT	2078			
DB	4116	GAAGTCCCTGTCGCTCTCTGTTCCGACCGCTGCGCTACCGGATACCTGTCCGCT	4175			

QY	2079	TTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGG	2138			
DB	4176	TTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGG	4235			
QY	2139	TGTAGGTGCTTCCGCTCAAGCTGGCTGTGTGACGAACCCCGCTTACGCCGACCGCT	2198			
DB	4236	TGTAGGTGCTTCCGCTCAAGCTGGCTGTGTGACGAACCCCGCTTACGCCGACCGCT	4295			
QY	2199	GGCGCTTATCCGGTAACCTATCGTCTTGAGTCCAAACCCGCTAAGACACGACTTTCGCCAC	2258			
DB	4296	GGCGCTTATCCGGTAACCTATCGTCTTGAGTCCAAACCCGCTAAGACACGACTTTCGCCAC	4355			
QY	2259	TGGCAGCAGCCACTGCTTAACAGGATTTAGCAGAGCGAGGTATGTAGGCGGTGCTCAGAGT	2318			
DB	4356	TGGCAGCAGCCACTGCTTAACAGGATTTAGCAGAGCGAGGTATGTAGGCGGTGCTCAGAGT	4415			
QY	2319	TCTTGAAGTGGTGGCTTAACCTACGGCTACACTAGAGGACAGTATTTGGTATCTCGGCTC	2378			
DB	4416	TCTTGAAGTGGTGGCTTAACCTACGGCTACACTAGAGGACAGTATTTGGTATCTCGGCTC	4475			
QY	2379	TGCTGAAGCCAGTTTACCTTCGGAAGAGAGTTGGTAGCTCTTGATCCGCGAAACAAACCA	2438			
DB	4476	TGCTGAAGCCAGTTTACCTTCGGAAGAGAGTTGGTAGCTCTTGATCCGCGAAACAAACCA	4535			
QY	2439	CCGCTGTAGCGGTGTTTTTTTGTTCGAAAGCAGCAGATTACGCGCAGAAAAAAGGAT	2498			
DB	4536	CCGCTGTAGCGGTGTTTTTTTGTTCGAAAGCAGCAGATTACGCGCAGAAAAAAGGAT	4595			
QY	2499	CTCAAGAGATCCTTTTGATCTTTTCTACGGGCTGTGACGCTCAGTGGGAACGAAACTCNC	2558			
DB	4596	CTCAAGAGATCCTTTTGATCTTTTCTACGGGCTGTGACGCTCAGTGGGAACGAAACTCNC	4655			
QY	2559	GTTAAGGATTTTGGTCAATGAGATTATCAAAAGAGTCTTCACTAGATCTCTTTTAAAT	2618			
DB	4656	GTTAAGGATTTTGGTCAATGAGATTATCAAAAGAGTCTTCACTAGATCTCTTTTAAAT	4715			
QY	2619	AAAAATGAAGTTTAAATCAATTAAGTATATATGAGTAAACTTTGGTCTGACAGTACC	2678			
DB	4716	AAAAATGAAGTTTAAATCAATTAAGTATATATGAGTAAACTTTGGTCTGACAGTACC	4775			
QY	2679	AATGCTTAATCAGTGAGGACCTATCTCAGGATCTGTCTATTTGTTTCATCCATAGTTG	2738			
DB	4776	AATGCTTAATCAGTGAGGACCTATCTCAGGATCTGTCTATTTGTTTCATCCATAGTTG	4835			
QY	2739	CTGACTCCCGTCTGTGTAGATACTACGATACGGAGGCTTACCATCTGCGCCCAAGTG	2798			
DB	4836	CTGACTCCCGTCTGTGTAGATACTACGATACGGAGGCTTACCATCTGCGCCCAAGTG	4895			
QY	2799	CTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTTATCAGCAATAAACACAGC	2858			
DB	4896	CTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTTATCAGCAATAAACACAGC	4955			
QY	2859	CAGCGGAAGGCGCGAGGAGAGTGTCTGTGAACTTTTATCCGCTCCATCCAGTCTA	2918			
DB	4956	CAGCGGAAGGCGCGAGGAGAGTGTCTGTGAACTTTTATCCGCTCCATCCAGTCTA	5015			
QY	2919	TTAATTGTTGCGGGAAGCTAGAGTAAGTGTTCGCCAGTTTAAATAGTTTGGCCAACTGTG	2978			
DB	5016	TTAATTGTTGCGGGAAGCTAGAGTAAGTGTTCGCCAGTTTAAATAGTTTGGCCAACTGTG	5075			
QY	2979	TTGCCATTGCTACAGGCACTGTGGTGTCAACGCTCGTGTGGTATAGCTTCAATTCAGCT	3038			
DB	5076	TTGCCATTGCTACAGGCACTGTGGTGTCAACGCTCGTGTGGTATAGCTTCAATTCAGCT	5135			
QY	3039	CGGTTTCCCAAGATCAAGGCGAGTTACATGATCCCGCATGTTGTGCAAAAGCGGTTA	3098			
DB	5136	CGGTTTCCCAAGATCAAGGCGAGTTACATGATCCCGCATGTTGTGCAAAAGCGGTTA	5195			
QY	3099	GCTCCTTCGCTCCCGATCGTTGTGCAAGTAAGTGTGCGCGCAGTGTATCACTCATGG	3158			
DB	5196	GCTCCTTCGCTCCCGATCGTTGTGCAAGTAAGTGTGCGCGCAGTGTATCACTCATGG	5255			
QY	3159	TTATGGCAGCACTGCATAAATTTCTCTTACTGTCTATGCCATCCGTAAGATGCTTTTCTGTGA	3218			

Db 5256 TTTATGCGAGCTGCATTAATTCCTTACTGTGTCATGCCATCCGTAAGATGCTTTTCTGTGA 5315
Qy 3219 CTGGTGAGTACTCAACCAAGTCACTTCTGAGAAATAGTGTATGCGGCGACCGAGTTGCTCTT 3278
Db 5316 CTGGTGAGTACTCAACCAAGTCACTTCTGAGAAATAGTGTATGCGGCGACCGAGTTGCTCTT 5375
Qy 3279 GCCCGCGCTCAATA CGGGATAATACCGCGCCACATAGCAGAACTTTAAAGTGTCTCATCA 3338
Db 5376 GCCCGCGCTCAATA CGGGATAATACCGCGCCACATAGCAGAACTTTAAAGTGTCTCATCA 5435
Qy 3339 TTGGAAAACGTTCTTCGGGGCGGAACCTCTCAAGGATCTTACCGCTGTGTGAGATCCAGTT 3398
Db 5436 TTGGAAAACGTTCTTCGGGGCGGAACCTCTCAAGGATCTTACCGCTGTGTGAGATCCAGTT 5495
Qy 3399 CGATGTAAACCCACTCGTGACCCCACTGATCTTTCAGCATCTTTTACTTTTACCAGCGTTT 3458
Db 5496 CGATGTAAACCCACTCGTGACCCCACTGATCTTTCAGCATCTTTTACTTTTACCAGCGTTT 5555
Qy 3459 CTGGGTGAGCAAAAACAGGAAGCAAAATGCCGCAAAAAGGAATAAGGGCGACACGGA 3518
Db 5556 CTGGGTGAGCAAAAACAGGAAGCAAAATGCCGCAAAAAGGAATAAGGGCGACACGGA 5615
Qy 3519 AATGTTGAATACCTCATACTCTTCTTCTTTTCAATATTATTGAAGCATTTATCAGGGTTATT 3578
Db 5616 AATGTTGAATACCTCATACTCTTCTTCTTTTCAATATTATTGAAGCATTTATCAGGGTTATT 5675
Qy 3579 GTCTCATGAGCGGATACATATTGTAATGATTTTAGAAAAATTAACAAATAGGGGTTCCGC 3638
Db 5676 GTCTCATGAGCGGATACATATTGTAATGATTTTAGAAAAATTAACAAATAGGGGTTCCGC 5735
Qy 3639 GCACATTTCCCGAAAAGTGCCAC 3662
Db 5736 GCACATTTCCCGAAAAGTGCCAC 5759

RESULT 14

US-09-404-066-21
; Sequence 21, Application US/09404066
; Patent No. 6365409
; GENERAL INFORMATION:
; APPLICANT: Bradley, John D.
; APPLICANT: Thompson, Craig M.
; APPLICANT: Moore, Jeffrey B.
; APPLICANT: Wobbe, C. Richard
; APPLICANT: Healy, Judith M.
; APPLICANT: Donnelly, Caroline E.
; TITLE OF INVENTION: REGULATED GENE EXPRESSION IN YEAST
; FILE REFERENCE: 0342/10469051
; CURRENT APPLICATION NUMBER: US/09/404,066
; CURRENT FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/138,024
; PRIOR FILING DATE: 1998-08-21
; PRIOR APPLICATION NUMBER: 60/056,719
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 21
; LENGTH: 7333
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Plasmid p2M197
US-09-404-066-21

Query Match 59.3%; Score 2172.8; DB 3; Length 7333;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2177; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
Qy 1479 GAGTTACCGCGCGCGCTTGGCGCTTAATCATGGTCATAGCTGTGTTCTGTTGAAATT 1538
Db 3576 GAGTCACCGCGATGCAAGCTTGGCGCTTAATCATGGTCATAGCTGTGTTCTGTTGAAATT 3635

Qy 1539 GTTATCCGCTCACAATTCACACAATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGG 1598
Db 3636 GTTATCCGCTCACAATTCACACAATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGG 3695
Qy 1599 GTGCTTAATGAGTGAGCTAACTCACAATTAATGCGTTGCGCTCAGCTGCCCGCTTTCCAGT 1658
Db 3696 GTGCTTAATGAGTGAGCTAACTCACAATTAATGCGTTGCGCTCAGCTGCCCGCTTTCCAGT 3755
Qy 1659 CGGGAAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAAACGCGCGGGAGAGCGGTT 1718
Db 3756 CGGGAAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAAACGCGCGGGAGAGCGGTT 3815
Qy 1719 TCGGTAATTTGGGCGCTCTTCCGCTTCTCGCTCAGCTGACCTCGCTGCGCTCGGTCGTTCCGC 1778
Db 3816 TCGGTAATTTGGGCGCTCTTCCGCTTCTCGCTCAGCTGACCTCGCTGCGCTCGGTCGTTCCGC 3875
Qy 1779 TCGGCGAGCGCTATCAGCTCACTCAAGCGCGTAAATACGGTTATCCACAGAAATCAAGGG 1838
Db 3876 TCGGCGAGCGCTATCAGCTCACTCAAGCGCGTAAATACGGTTATCCACAGAAATCAAGGG 3935
Qy 1839 ATAAACGAGGAAGAAACATGTGAGCAAAAGGCGCAAAAGGCGCAGAAACCGTAAAAAGG 1898
Db 3936 ATAAACGAGGAAGAAACATGTGAGCAAAAGGCGCAGCAAAAGGCGCAGAAACCGTAAAAAGG 3995
Qy 1899 CGCGCTGTCGCGCTTTTCCATAGGCTCGGCCCTCGCTGACGAGCATCACAAAATCGAC 1958
Db 3996 CGCGCTGTCGCGCTTTTCCATAGGCTCGGCCCTCGCTGACGAGCATCACAAAATCGAC 4055
Qy 1959 GCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAAGGCGTTTCCCGCTG 2018
Db 4056 GCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAAGGCGTTTCCCGCTG 4115
Qy 2019 GAAGCTCCCTCGTGCCTCTCTCTGTTCCGACCCCTGCGCTTACCGGATACCTGTCCGCT 2078
Db 4116 GAAGCTCCCTCGTGCCTCTCTCTGTTCCGACCCCTGCGCTTACCGGATACCTGTCCGCT 4175
Qy 2079 TTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCCG 2138
Db 4176 TTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCCG 4235
Qy 2139 TGTAGTGTCTCGCTCCAAAGCTGGGCTGTGCAGCAACCCCGCTTACGCGCGACCGCT 2198
Db 4236 TGTAGTGTCTCGCTCCAAAGCTGGGCTGTGCAGCAACCCCGCTTACGCGCGACCGCT 4295
Qy 2199 GCGCTTATCCGCTAACTATCGTCTTGTAGTCCAAACCCCGTAAACGACGACTTATCGCCAC 2258
Db 4296 GCGCTTATCCGCTAACTATCGTCTTGTAGTCCAAACCCCGTAAACGACGACTTATCGCCAC 4355
Qy 2259 TGGCAGAGCCACTGGTAAACAGGATTTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAT 2318
Db 4356 TGGCAGAGCCACTGGTAAACAGGATTTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAT 4415
Qy 2319 TCTTGAAGTGGTGGCTTAACGCTACACGCTACAGAGGACAGTATTGTGTATCTCGCTC 2378
Db 4416 TCTTGAAGTGGTGGCTTAACGCTACACGCTACAGAGGACAGTATTGTGTATCTCGCTC 4475
Qy 2379 TGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTTCATCCGCCAAACCA 2438
Db 4476 TGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTTCATCCGCCAAACCA 4535
Qy 2439 CGCGTGTAGCGGTGTTTTTTTTTGTTCGAAAGCAGATTTAGCGCGAGAAAAAGGAT 2498
Db 4536 CGCGTGTAGCGGTGTTTTTTTTTGTTCGAAAGCAGATTTAGCGCGAGAAAAAGGAT 4595
Qy 2499 CTCAGAAGATCCTTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAAACGAAACTCAC 2558
Db 4596 CTCAGAAGATCCTTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAAACGAAACTCAC 4655
Qy 2559 GTTAAAGGATTTTGGTTCATGAGATTATCAAAAGGATCTTTCACCTAGATCTTTTAAAT 2618
Db 4656 GTTAAAGGATTTTGGTTCATGAGATTATCAAAAGGATCTTTCACCTAGATCTTTTAAAT 4715
Qy 2619 AAAAATGAAGTTTTAAATCAATCAATGAAGTATATATGAGTAAACTTGGTCTGACAGTTACC 2678

|||||
4716 AAAAAATGAAGTATTAATCAATCTAAAGTATATATAGTAAAGCTTGGTCTGACAGTTACC 4775
QY AATGCTTAATCAGTGGAGGACCATCTCTCAGCGATCTGTCTATTTCTGTTCCATCCATAGTTG 2738
Db AATGCTTAATCAGTGGAGGACCATCTCTCAGCGATCTGTCTATTTCTGTTCCATCCATAGTTG 4835
QY CTTGACTCTCCCGTCTGTGTAGATAAATACATACGGGAGGCTTACCATCTGGCCCCAGTG 2798
Db CTTGACTCTCCCGTCTGTGTAGATAAATACATACGGGAGGCTTACCATCTGGCCCCAGTG 4895
QY CTGCAATGATACCGGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGC 2858
Db CTGCAATGATACCGGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGC 4955
QY CAGCCGGAAGGCGGAGCGCAGAGTGGTCTGCAACTTTATCCGCTCCATCCAGTCTA 2918
Db CAGCCGGAAGGCGGAGCGCAGAGTGGTCTGCAACTTTATCCGCTCCATCCAGTCTA 5015
QY TTAATTTGTTGCCGGGAAGCTAGAGTAAGTATTCGCCAGTAAATAGTTTGGCCAAAGTTG 2978
Db TTAATTTGTTGCCGGGAAGCTAGAGTAAGTATTCGCCAGTAAATAGTTTGGCCAAAGTTG 5075
QY TTGCCATTGCTACAGCATCTGTGGTGTACGCTCGTGGTTGGTATGGCTTCATTCAGCT 3038
Db TTGCCATTGCTACAGCATCTGTGGTGTACGCTCGTGGTATGGCTTCATTCAGCT 5135
QY CGGTTTCCCAACGATCAGGCGAGTACATGATCCCCATGTTGTGCAAAAAGCGGTTA 3098
Db CGGTTTCCCAACGATCAGGCGAGTACATGATCCCCATGTTGTGCAAAAAGCGGTTA 5195
QY GCTCTCTCGGTCCTCCGATCTGTTGAGAGTAAGTTGGCCGAGTGTATCACTCATCG 3158
Db GCTCTCTCGGTCCTCCGATCTGTTGAGAGTAAGTTGGCCGAGTGTATCACTCATCG 5255
QY TTATGGCAGCATGCATTAATCTCTTACTGTGATGCTATGCTGCTTCTTCTGTTG 3218
Db TTATGGCAGCATGCATTAATCTCTTACTGTGATGCTATGCTGCTTCTTCTGTTG 5315
QY CTGGTGGTACTCAACCAAGTCAATCTGAGATAAGTATGCGGCGACCGAGTTGCTCT 3278
Db CTGGTGGTACTCAACCAAGTCAATCTGAGATAAGTATGCGGCGACCGAGTTGCTCT 5375
QY GCCCGGCTCAATACGGGATAATACCGGCCACATAGCAGAACTTTAAAGTGTCTATCA 3338
Db GCCCGGCTCAATACGGGATAATACCGGCCACATAGCAGAACTTTAAAGTGTCTATCA 5435
QY TTGGAAGAGCTTCTTGGGGCGAAATCTCTCAAGATCTTACCGTGTGTGAGATCCAGTT 3398
Db TTGGAAGAGCTTCTTGGGGCGAAATCTCTCAAGATCTTACCGTGTGTGAGATCCAGTT 5495
QY CGATGTAACCCACTCGTGACCCCACTGATCTTACGATCTTTTACTTTTCAACGCTTT 3458
Db CGATGTAACCCACTCGTGACCCCACTGATCTTACGATCTTTTACTTTTCAACGCTTT 5555
QY CTGGGTGAGCAAAAACAGGAAGCAAAATGCCGCAAAAAGGGAATAAGGGCGACACGGA 3518
Db CTGGGTGAGCAAAAACAGGAAGCAAAATGCCGCAAAAAGGGAATAAGGGCGACACGGA 5615
QY AATGTTGAATATCTCATATCTTCTCTTTTCAATATTTATGAGCAATTTATCAGGTTAT 3578
Db AATGTTGAATATCTCATATCTTCTCTTTTCAATATTTATGAGCAATTTATCAGGTTAT 5675
QY GTCTCATGCGGATACATATTTGATGTTATTTAGAAAATAAACAATAGGGTTCCGC 3638
Db GTCTCATGCGGATACATATTTGATGTTATTTAGAAAATAAACAATAGGGTTCCGC 5735
QY GCACATTTCCCGAAAAGTGCCAC 3662
Db GCACATTTCCCGAAAAGTGCCAC 5759

RESULT 15

US-09-573-322-21
; Sequence 21, Application US/09573322
; Patent No. 6531289
; GENERAL INFORMATION:
; APPLICANT: Bradley, John D.
; APPLICANT: Thompson, Craig M.
; APPLICANT: Moore, Jeffrey B.
; APPLICANT: Wobbe, C. Richard
; APPLICANT: Bailey, David A.
; TITLE OF INVENTION: Regulated Gene Expression in Yeast and
; TITLE OF INVENTION: Methods of Use
; FILE REFERENCE: 0342/ID469-US4
; CURRENT APPLICATION NUMBER: US/09/573,322
; CURRENT FILING DATE: 2000-05-18
; PRIOR APPLICATION NUMBER: 09/404,066
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 09/138,024
; PRIOR FILING DATE: 1998-08-21
; PRIOR APPLICATION NUMBER: 60/056,719
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 21
; LENGTH: 7333
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: ZM197 plasmid
US-09-573-322-21

Query Match 59.3%; Score 2172.8; DB 4; Length 7333;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2177; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1479 GAGTTACCGCGCGCGCTTGGGTAATCATGTGTATAGTGTCTTCTGTGTAATTT 1538
Db 3576 GAGTCGACGGCATCAAGCTTGGGTAATCATGTGTATAGTGTCTTCTGTGTAATTT 3635
QY 1539 GTTATCCGCTCACATTTCCACACATACGAGCGGAAGCATAAAGTGTAAAGCTGGG 1598
Db 3636 GTTATCCGCTCACATTTCCACACATACGAGCGGAAGCATAAAGTGTAAAGCTGGG 3695
QY 1599 GTGCCATAATGAGTGAAGTCACTCAATTAATTTGCGTTGCGTCACTGCGCTTTCCAGT 1658
Db 3696 GTGCCATAATGAGTGAAGTCACTCAATTAATTTGCGTTGCGTCACTGCGCTTTCCAGT 3755
QY 1659 CGGGAACCTGTCGTCAGCTGCATTAATGAATCGGCAACGCGCGGAGAGCGGTT 1718
Db 3756 CGGGAACCTGTCGTCAGCTGCATTAATGAATCGGCAACGCGCGGAGAGCGGTT 3815
QY 1719 TGCGTATTTGGCGCTCTTCCGCTTCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCCGC 1778
Db 3816 TGCGTATTTGGCGCTCTTCCGCTTCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCCGC 3875
QY 1779 TGCGCGAGCGGTATCAGTCACTCAAAAGCGGTAATACGGTTATCCACAGAAATCAGGG 1838
Db 3876 TGCGCGAGCGGTATCAGTCACTCAAAAGCGGTAATACGGTTATCCACAGAAATCAGGG 3935
QY 1839 ATAACGAGGAAAGAAATGATGAGCAAAAGCCAGCAAAAGCCAGGACCGTAAAGG 1898
Db 3936 ATAACGAGGAAAGAAATGATGAGCAAAAGCCAGCAAAAGCCAGGACCGTAAAGG 3995
QY 1899 CCGGTTCTCGCGCTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCCAC 1958
Db 3996 CCGGTTCTCGCGCTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCCAC 4055
QY 1959 GCTCAAGTCAGAGGTGGCGAAAACCGACAGGACTATAAAGATACAGCGGTTTCCCCCTG 2018
Db 4056 GCTCAAGTCAGAGGTGGCGAAAACCGACAGGACTATAAAGATACAGCGGTTTCCCCCTG 4115
QY 2019 GAAGCTCCCTCGCGCTCTCTGTTCCGACCTGCGCTTACCGGATACCTGTCCGCT 2078
Db 4116 GAAGCTCCCTCGCGCTCTCTGTTCCGACCTGCGCTTACCGGATACCTGTCCGCT 4175

QY	2079	TTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCCG	2138
Db	4176	TTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCCG	4235
QY	2139	TGTAGGTCTGCTCCAAAGCTGGGCTGTGTGCAGAAACCCCGCTTACAGCCGACCGCT	2198
Db	4236	TGTAGGTCTGCTCCAAAGCTGGGCTGTGTGCAGAAACCCCGCTTACAGCCGACCGCT	4295
QY	2199	GGCCCTTATCCGGTAACTATCGTTTGGAGTCCAAACCCCGTAAAGACACGACTTATCCGCAC	2258
Db	4296	GGCCCTTATCCGGTAACTATCGTTTGGAGTCCAAACCCCGTAAAGACACGACTTATCCGCAC	4355
QY	2259	TGGCAGCAGCCACTGTATACAGGATTTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAT	2318
Db	4356	TGGCAGCAGCCACTGTATACAGGATTTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAT	4415
QY	2319	TCTTGAAGTGGTGGCTTAACCTACCGCTACACTAGAAAGGACAGTATTTGGTATCTGGCTC	2378
Db	4416	TCTTGAAGTGGTGGCTTAACCTACCGCTACACTAGAAAGGACAGTATTTGGTATCTGGCTC	4475
QY	2379	TGCTGAAGCCAGTTACCTTCGGAAGAGAGTTGGTAGCTCTTGATCCGGCAACCAACCA	2438
Db	4476	TGCTGAAGCCAGTTACCTTCGGAAGAGAGTTGGTAGCTCTTGATCCGGCAACCAACCA	4535
QY	2439	CGCTGGTAGCGGTGGTTTTTTTGGTGAAGCAGAGATTAGCGCGCAGAAAGGAT	2498
Db	4536	CGCTGGTAGCGGTGGTTTTTTTGGTGAAGCAGAGATTAGCGCGCAGAAAGGAT	4595
QY	2499	CTCAAGAGACTCTTTGATCTTTCTACCGGGCTCTGACCGCTCAGTGGAAAGAACTCAC	2558
Db	4596	CTCAAGAGACTCTTTGATCTTTCTACCGGGCTCTGACCGCTCAGTGGAAAGAACTCAC	4655
QY	2559	GTTAAGGATTTTGGTTCATGAGATTATCAAAAGGATCTTACCTAGATCTTTTAAAT	2618
Db	4656	GTTAAGGATTTTGGTTCATGAGATTATCAAAAGGATCTTACCTAGATCTTTTAAAT	4715
QY	2619	AAAAATGAAGTTTAAATCAATCTAAAGTATATATAGTAAACTTGGTCTGACAGTTACC	2678
Db	4716	AAAAATGAAGTTTAAATCAATCTAAAGTATATATAGTAAACTTGGTCTGACAGTTACC	4775
QY	2679	AATGCTTAATCAGTGAGGACCTATCTCAGCGATCTGTCTATTTGTTTCATCCATAGTTG	2738
Db	4776	AATGCTTAATCAGTGAGGACCTATCTCAGCGATCTGTCTATTTGTTTCATCCATAGTTG	4835
QY	2739	CCTGACTCCCGCTGTGTAGATAACTACGATACGGGAGGCTTACCATCTGGCCCCAGTG	2798
Db	4836	CCTGACTCCCGCTGTGTAGATAACTACGATACGGGAGGCTTACCATCTGGCCCCAGTG	4895
QY	2799	CTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTATCAGCAATAAACCCAGC	2858
Db	4896	CTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTATCAGCAATAAACCCAGC	4955
QY	2859	CAGCCGGAAGCGCGAGCGCAGAGTGGTCTGCAACTTTATCCGCTCCATCCAGTCTA	2918
Db	4956	CAGCCGGAAGCGCGAGCGCAGAGTGGTCTGCAACTTTATCCGCTCCATCCAGTCTA	5015
QY	2919	TTAATTTGTTGGCGGAAGCTAGAGTAAAGTATTCGCCAGTAAATAGTTTGGCGCAAGTTG	2978
Db	5016	TTAATTTGTTGGCGGAAGCTAGAGTAAAGTATTCGCCAGTAAATAGTTTGGCGCAAGTTG	5075
QY	2979	TTGCCATTGCTACAGGCATCGTGGTGTCA CGCTCGTGGTTGGTATGGCTTCATTAGCT	3038
Db	5076	TTGCCATTGCTACAGGCATCGTGGTGTCA CGCTCGTGGTTGGTATGGCTTCATTAGCT	5135
QY	3039	CCGGTTCCTCAACGATCAAGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTA	3098
Db	5136	CCGGTTCCTCAACGATCAAGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTA	5195
QY	3099	GCTCTTCGGTCTCTCCGATCGTTGTGAGAAAGTAAAGTTGGCGCAGTGGTTATCACTCATGG	3158
Db	5196	GCTCTTCGGTCTCTCCGATCGTTGTGAGAAAGTAAAGTTGGCGCAGTGGTTATCACTCATGG	5255

QY	3159	TTATGGCAGCACTGCATATAATTTCTTACTGTGTATGCCATCCGTAAAGATGCTTTTCTGTGA	3218
Db	5256	TTATGGCAGCACTGCATATAATTTCTTACTGTGTATGCCATCCGTAAAGATGCTTTTCTGTGA	5315
QY	3219	CTGGTGAGTACTCAACCAAGTCAATCTGAGAAATAGTGTATGCGGCGACCGAGTTGCTCTT	3278
Db	5316	CTGGTGAGTACTCAACCAAGTCAATCTGAGAAATAGTGTATGCGGCGACCGAGTTGCTCTT	5375
QY	3279	GCCCGGCGTCAATA CGGGATAATACCGGCCACATAGCAGAACTTTTAAAGTCTCATCA	3338
Db	5376	GCCCGGCGTCAATA CGGGATAATACCGGCCACATAGCAGAACTTTTAAAGTCTCATCA	5435
QY	3339	TTTGGAAAAAGTTCTTTCGGGGCGAAGAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTT	3398
Db	5436	TTTGGAAAAAGTTCTTTCGGGGCGAAGAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTT	5495
QY	3399	CGATGTAAACCCCACTCGTGACCCCACTGATCTTTCAGCATCTTTTACTTTTACCAGCGTTT	3458
Db	5496	CGATGTAAACCCCACTCGTGACCCCACTGATCTTTCAGCATCTTTTACTTTTACCAGCGTTT	5555
QY	3459	CTGGGTGAGCAAAAAACAGGAAGCAAAATGCGCGCAAAAAAGGAAATAAGGGCGACACGGA	3518
Db	5556	CTGGGTGAGCAAAAAACAGGAAGCAAAATGCGCGCAAAAAAGGAAATAAGGGCGACACGGA	5615
QY	3519	AATGTTGAATACTCATACTCTTCTTTTCAATATTTTGAAGCATTTATCAGGGTTATT	3578
Db	5616	AATGTTGAATACTCATACTCTTCTTTTCAATATTTTGAAGCATTTATCAGGGTTATT	5675
QY	3579	GTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAACAAATAGGGGTTCCGC	3638
Db	5676	GTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAACAAATAGGGGTTCCGC	5735
QY	3639	GCACATTTTCCCGAAAAAGTGCCAC	3662
Db	5736	GCACATTTTCCCGAAAAAGTGCCAC	5759

Search completed: June 5, 2005, 23:08:18
Job time : 656.194 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 5, 2005, 13:38:04 ; Search time 10710.2 Seconds
(without alignments)
13014.775 Million cell updates/sec

Title: US-10-826-523-41

Perfect score: 3662

Sequence: 1 ctaaatgtaagcgttaata.....attccccgaaagtgcacc 3662

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

- 1: gb_est1:*
- 2: gb_est2:*
- 3: gb_hic:*
- 4: gb_est3:*
- 5: gb_est4:*
- 6: gb_est5:*
- 7: gb_est6:*
- 8: gb_gse1:*
- 9: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1055.8	28.8	1070	1	AJ281552 4A3A-P6F1
2	1018.2	27.8	1048	7	CO552396 AC14 50
3	986.4	26.9	1013	4	BM438846 IPLvr0015
4	955.2	26.1	1067	1	AU081137 AU081137
5	924	25.2	1004	1	AJ281480 4A3A-P4G8
6	923.8	25.2	1089	1	AU081124 AU081124
7	902.8	24.7	1049	9	CL021189 CH216-8A1
8	889.8	24.3	928	7	CO487414 GQ0227.B7
9	881.4	24.1	925	6	CB686151 Bn01b_020
10	876.4	23.9	902	7	CR753463 DKFZp469G
11	870.2	23.8	935	4	BG838279 GC01_10e0
12	865	23.6	1073	7	CF269652 FcylcolD8
13	857.6	23.4	885	9	CL076016 CH216-138
14	841	23.0	841	1	AL042026 DKFZp434E
15	824.6	22.5	854	4	BM438950 IPLvr0049
16	824	22.5	1169	9	AG332951 Mus muscu
17	821.4	22.4	1025	9	CL021193 CH216-8A1
18	820.2	22.4	865	7	CK125894 BES182411
19	819.6	22.4	1249	8	BZ572284 msh2_2572
20	814.4	22.2	1126	8	BZ577702 msh2_5533
21	809.4	22.1	856	7	CR823189 Oa splbn
22	808	22.1	819	6	CD649375 CVGn00008
23	807.6	22.1	966	8	BZ570738 msh2_1513
24	797.6	21.8	1011	8	BZ576726 msh2_5071

ALIGNMENTS

RESULT 1

LOCUS	AJ281552	1070 bp	mRNA	linear	EST 30-JUN-2000
DEFINITION	4A3A-P6F11-F Anopheles gambiae immune competent 4A3A Anopheles gambiae cDNA clone 4A3A-P6F11, mRNA sequence.				
ACCESSION	AJ281552				
VERSION	AJ281552.1	GI:6929432			
KEYWORDS	EST.				
SOURCE	Anopheles gambiae (African malaria mosquito)				
ORGANISM	Anopheles gambiae				
REFERENCE	1 (bases 1 to 1070)				
AUTHORS	Dimopoulos, G., Csaevant, T.L., Chang, S., Scheetz, T., Roberts, C., Donohue, M., Schultz, J., Benes, V., Bork, P., Ansoorge, W., Soares, M.B. and Kafatos, F.C.				
TITLE	Anopheles gambiae pilot gene discovery project: identification of mosquito innate immunity genes from expressed sequence tags generated from immune-competent cell lines				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)				
MEDLINE	20300950				
PUBMED	10841561				
COMMENT	Contact: Dimopoulos G Potis C. Kafatos laboratory European Molecular Biology Laboratory Meyerhofstrasse 1, 69117 Heidelberg, Germany. Location/Qualifiers i. .1070 /organism="Anopheles gambiae" /mol_type="mRNA" /strain="4A r/r" /db_xref="taxon:7165" /clone="4A3A-P6F11" /lab_host="E. coli DH10B" /note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: EcoRI; Site 2: NotI; sequenced from forward priming site which reads from the 3' end of the cDNA. The 4A3A is a directionally cloned and normalized cDNA library that was constructed from the 4A3A cell line oligo-T primed cDNA according to: Bonaldo, Lennon & Soares (1996) - Normalization and Subtraction: Two approaches To Facilitate Gene Discovery, Genome Research 6, 791-806."				
FEATURES	source				
ORIGIN					

Query Match		28.8%; Score 1055.8; DB 1; Length 1070;	
Best Local Similarity		99.7%; Pred. No. 3.9e-299;	
Matches 1068; Conservative		0; Mismatches 2; Indels 1; Gaps 1;	
QY	1904	TTGCTGGCGT	TTTTCCATAGGCTCGCCCCCTTCGAGGACATCACAAAATCGAGCTCA 1963
DB	1	TTGCTGGCGT	TTTTCCATAGGCTCGCCCCCTTCGAGGACATCACAAAATCGAGCTCA 60
QY	1964	AGTCAGAGGTGGCGAAACCCGACAGACTATAAAGATACAGGCGT	TTTCCCTCTGGAGC 2023
DB	61	AGTCAGAGGTGGCGAAACCCGACAGACTATAAAGATACAGGCGT	TTTCCCTCTGGAGC 120
QY	2024	TCCCTCGTGGCTCTCTGTTCCGACCCCTGCGCTTACCGGATACCTGTCGCGCTTTCTC 2083	
DB	121	TCCCTCGTGGCTCTCTGTTCCGACCCCTGCGCTTACCGGATACCTGTCGCGCTTTCTC 180	
QY	2084	CCTTCGGGAAGCGTGGCGCTTCTCATAGCTCACGCTGTAGTATCTCAGTTTGGGTAG 2143	
DB	181	CCTTCGGGAAGCGTGGCGCTTCTCATAGCTCACGCTGTAGTATCTCAGTTTGGGTAG 240	
QY	2144	GTTCGTTCCGCTCAAGCTGGGCTGTGTGCACGAACCCCGCTTACGCCGACCGCTGCGCC 2203	
DB	241	GTTCGTTCCGCTCAAGCTGGGCTGTGTGCACGAACCCCGCTTACGCCGACCGCTGCGCC 300	
QY	2204	TTATCCGGTAACTATCGTCTTGAGTCCAAACCCGCTTAAGACACGACTTATCGCACTGGCA 2263	
DB	301	TTATCCGGTAACTATCGTCTTGAGTCCAAACCCGCTTAAGACACGACTTATCGCACTGGCA 360	
QY	2264	CGAGCACTGGTAAACAGATTTAGCAGAGGAGGTATGTAGGCGGTCTACAGATTTCTTG 2323	
DB	361	CGAGCACTGGTAAACAGATTTAGCAGAGGAGGTATGTAGGCGGTCTACAGATTTCTTG 420	
QY	2324	AAGTGGTGGCTTAACTACGGCTTACACTAGAGGACAGTATTTGGTATCTGGCTCTGGT 2383	
DB	421	AAGTGGTGGCTTAACTACGGCTTACACTAGAGGACAGTATTTGGTATCTGGCTCTGGT 480	
QY	2384	AAGCCAGTTTACCTTCGGAAGAGTGGTGTAGTCTTGTATCGGCAAAACAAACACCGCT 2443	
DB	481	AAGCCAGTTTACCTTCGGAAGAGTGGTGTAGTCTTGTATCGGCAAAACAAACACCGCT 540	
QY	2444	GGTAGCGTGGTTTTTTTGGTTGGAAGCAGAGATTAACCGCAGAAAAAAGGATCTCAA 2503	
DB	541	GGTAGCGTGGTTTTTTTGGTTGGAAGCAGAGATTAACCGCAGAAAAAAGGATCTCAA 600	
QY	2504	GAAAGATCTTTGATCTTTCTACGGGCTCTAGCGCTCAGTGGAGCAAAACCTCAGCTTAA 2563	
DB	601	GAAAGATCTTTGATCTTTCTACGGGCTCTAGCGCTCAGTGGAGCAAAACCTCAGCTTAA 660	
QY	2564	GGGATTTTGGTCATGAGATTTATCAAAAAGGATCTTCACTAGATCTTTTAAATTAATA 2623	
DB	661	GGGATTTTGGTCATGAGATTTATCAAAAAGGATCTTCACTAGATCTTTTAAATTAATA 720	
QY	2624	TGAAGTTTAAATCAATCTAAGTATATATAGTAAACTTTGGTCTGACAGTTTACCAATGC 2683	
DB	721	TGAAGTTTAAATCAATCTAAGTATATATAGTAAACTTTGGTCTGACAGTTTACCAATGC 780	
QY	2684	TTAATCAGTAGGCACTTCTCAGGATCTGTCTATTTTGGTTTCATCCATAGTTCGCTGA 2743	
DB	781	TTAATCAGTAGGCACTTCTCAGGATCTGTCTATTTTGGTTTCATCCATAGTTCGCTGA 840	
QY	2744	CTCCCGCTGCTGTAGATAACTACGATACGGAGGGCTTACCATCTGGCCCCCAGTCTGCA 2803	
DB	841	CTCCCGCTGCTGTAGATAACTACGATACGGAGGGCTTACCATCTGGCCCCCAGTCTGCA 900	
QY	2804	ATGATACCGCAGAGACCGCTCACCGGCTCCAGATTTTATCAGCAATTAACACGACGCC 2863	
DB	901	ATGATACCGCAGAGACCGCTCACCGGCTCCAGATTTTATCAGCAATTAACACGACGCC 960	
QY	2864	GGAAGGCGGACGCGAGAGGTGCTCTGCACTTTTATCCGCTCCATCCAGTCTATTAT 2923	
DB	961	GGAA--GGCGGACGCGAGAGGTGCTCTGCACTTTTATCCGCTCCATCCAGTCTATTAT 1019	
QY	2924	TGTTGCCGGGAAGCTAGAGTAAGTATGTTTCGCCAGTTAATAGTTTTCGCCAAC 2974	
DB	1020	TGTTCCCGGAAGCTAGAGTAAGTATGTTTCGCCAGTTAATAGTTTTCGCCAAC 1070	
RESULT 2			
COS52396			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			
1..1048			
/organism="Petromyzon marinus"			
/mol_type="mRNA"			
/db_xref="taxon:7757"			
/cell_type="lymphocyte"			
/dev_stage="immune stimulated larvae"			
/clone_lib="Sea lamprey Acly"			
/note="Vector: pGEM-T Easy; lymphocyte mRNA ESTs from PCR			
subtracted cDNA libraries of immune stimulated larvae. All			
are single pass 5' or 3' sequences randomly cloned in			
pGEM-T Easy (Promega)."			
Location/Qualifiers			
1..1048			
/organism="Petromyzon marinus"			
/mol_type="mRNA"			
/db_xref="taxon:7757"			
/cell_type="lymphocyte"			
/dev_stage="immune stimulated larvae"			
/clone_lib="Sea lamprey Acly"			
/note="Vector: pGEM-T Easy; lymphocyte mRNA ESTs from PCR			
subtracted cDNA libraries of immune stimulated larvae. All			
are single pass 5' or 3' sequences randomly cloned in			
pGEM-T Easy (Promega)."			
Location/Qualifiers			
1..1048			
/organism="Petromyzon marinus"			
/mol_type="mRNA"			
/db_xref="taxon:7757"			
/cell_type="lymphocyte"			
/dev_stage="immune stimulated larvae"			
/clone_lib="Sea lamprey Acly"			
/note="Vector: pGEM-T Easy; lymphocyte mRNA ESTs from PCR			
subtracted cDNA libraries of immune stimulated larvae. All			
are single pass 5' or 3' sequences randomly cloned in			
pGEM-T Easy (Promega)."			
Location/Qualifiers			
1..1048			
/organism="Petromyzon marinus"			
/mol_type="mRNA"			
/db_xref="taxon:7757"			
/cell_type="lymphocyte"			
/dev_stage="immune stimulated larvae"			
/clone_lib="Sea lamprey Acly"			
/note="Vector: pGEM-T Easy; lymphocyte mRNA ESTs from PCR			
subtracted cDNA libraries of immune stimulated larvae. All			
are single pass 5' or 3' sequences randomly cloned in			
pGEM-T Easy (Promega)."			
Location/Qualifiers			
1..1048			
/organism="Petromyzon marinus"			
/mol_type="mRNA"			
/db_xref="taxon:7757"			
/cell_type="lymphocyte"			
/dev_stage="immune stimulated larvae"			
/clone_lib="Sea lamprey Acly"			
/note="Vector: pGEM-T Easy; lymphocyte mRNA ESTs from PCR			
subtracted cDNA libraries of immune stimulated larvae. All			
are single pass 5' or 3' sequences randomly cloned in			
pGEM-T Easy (Promega)."			
Location/Qualifiers			
1..1048			
/organism="Petromyzon marinus"			
/mol_type="mRNA"			
/db_xref="taxon:7757"			
/cell_type="lymphocyte"			
/dev_stage="immune stimulated larvae"			
/clone_lib="Sea lamprey Acly"			
/note="Vector: pGEM-T Easy; lymphocyte mRNA ESTs from PCR			
subtracted cDNA libraries of immune stimulated larvae. All			
are single pass 5' or 3' sequences randomly cloned in			
pGEM-T Easy (Promega)."			
Location/Qualifiers			
1..1048			
/organism="Petromyzon marinus"			
/mol_type="mRNA"			
/db_xref="taxon:7757"			
/cell_type="lymphocyte"			
/dev_stage="immune stimulated larvae"			
/clone_lib="Sea lamprey Acly"			
/note="Vector: pGEM-T Easy; lymphocyte mRNA ESTs from PCR			
subtracted cDNA libraries of immune stimulated larvae. All			
are single pass 5' or 3' sequences randomly cloned in			
pGEM-T Easy (Promega)."			
Location/Qualifiers			
1..1048			
/organism="Petromyzon marinus"			
/mol_type="mRNA"			
/db_xref="taxon:7757"			
/cell_type="lymphocyte"			
/dev_stage="immune stimulated larvae"			
/clone_lib="Sea lamprey Acly"			
/note="Vector: pGEM-T Easy; lymphocyte mRNA ESTs from PCR			
subtracted cDNA libraries of immune stimulated larvae. All			
are single pass 5' or 3' sequences randomly cloned in			
pGEM-T Easy (Promega)."			
Location/Qualifiers			
1..1048			
/organism="Petromyzon marinus"			
/mol_type="mRNA"			
/db_xref="taxon:7757"			
/cell_type="lymphocyte"			
/dev_stage="immune stimulated larvae"			
/clone_lib="Sea lamprey Acly"			
/note="Vector: pGEM-T Easy; lymphocyte mRNA ESTs from PCR			
subtracted cDNA libraries of immune stimulated larvae. All			
are single pass 5' or 3' sequences randomly cloned in			
pGEM-T Easy (Promega)."			
Location/Qualifiers			
1..1048			
/organism="Petromyzon marinus"			
/mol_type="mRNA"			
/db_xref="taxon:7757"			
/cell_type="lymphocyte"			
/dev_stage="immune stimulated larvae"			
/clone_lib="Sea lamprey Acly"			
/note="Vector: pGEM-T Easy; lymphocyte mRNA ESTs from PCR			
subtracted cDNA libraries of immune stimulated larvae. All			
are single pass 5' or 3' sequences randomly cloned in			
pGEM-T Easy (Promega)."			
Location/Qualifiers			
1..1048			
/organism="Petromyzon marinus"			
/mol_type="mRNA"			
/db_xref="taxon:7757"			
/cell_type="lymphocyte"			
/dev_stage="immune stimulated larvae"			
/clone_lib="Sea lamprey Acly"			
/note="Vector: pGEM-T Easy; lymphocyte mRNA ESTs from PCR			
subtracted cDNA libraries of immune stimulated larvae. All			
are single pass 5' or 3' sequences randomly cloned in			
pGEM-T Easy (Promega)."			
Location/Qualifiers			
1..1048			
/organism="Petromyzon marinus"			
/mol_type="mRNA"			
/db_xref="taxon:7757"			
/cell_type="lymphocyte"			
/dev_stage="immune stimulated larvae"			
/clone_lib="Sea lamprey Acly"			
/note="Vector: pGEM-T Easy; lymphocyte mRNA ESTs from PCR			
subtracted cDNA libraries of immune stimulated larvae. All			
are single pass 5' or 3' sequences randomly cloned in			
pGEM-T Easy (Promega)."			
Location/Qualifiers			
1..1048			
/organism="Petromyzon marinus"			
/mol_type="mRNA"			
/db_xref="taxon:7757"			
/cell_type="lymphocyte"			
/dev_stage="immune stimulated larvae"			
/clone_lib="Sea lamprey Acly"			
/note="Vector: pGEM-T Easy; lymphocyte mRNA ESTs from PCR			
subtracted cDNA libraries of immune stimulated larvae. All			
are single pass 5' or 3' sequences randomly cloned in			
pGEM-T Easy (Promega)."			
Location/Qualifiers			
1..1048			
/organism="Petromyzon marinus"			
/mol_type="mRNA"			
/db_xref="taxon:7757"			
/cell_type="lymphocyte"			
/dev_stage="immune stimulated larvae"			
/clone_lib="Sea lamprey Acly"			
/note="Vector: pGEM-T Easy; lymphocyte mRNA ESTs from PCR			
subtracted cDNA libraries of immune stimulated larvae. All			
are single pass 5' or 3' sequences randomly cloned in			
pGEM-T Easy (Promega)."			
Location/Qualifiers			
1..1048			
/organism="Petromyzon marinus"			
/mol_type="mRNA"			
/db_xref="taxon:7757"			
/cell_type="lymphocyte"			
/dev_stage="immune stimulated larvae"			
/clone_lib="Sea lamprey Acly"			
/note="Vector: pGEM-T Easy; lymphocyte mRNA ESTs from PCR			
subtracted cDNA libraries of immune stimulated larvae. All			
are single pass 5' or 3' sequences randomly cloned in			
pGEM-T Easy (Promega)."			
Location/Qualifiers			
1..1048			
/organism="Petromyzon marinus"			
/mol_type="mRNA"			
/db_xref="taxon:7757"			
/cell_type="lymphocyte"			
/dev_stage="immune stimulated larvae"			
/clone_lib="Sea lamprey Acly"			
/note="Vector: pGEM-T Easy; lymphocyte mRNA ESTs from PCR			
subtracted cDNA libraries of immune stimulated larvae. All			
are single pass 5' or 3' sequences randomly cloned in			
pGEM-T Easy (Promega)."			
Location/Qualifiers			
1..1048			
/organism="Petromyzon marinus"			
/mol_type="mRNA"			
/db_xref="taxon:7757"			
/cell_type="lymphocyte"			
/dev_stage="immune stimulated larvae"			
/clone_lib="Sea lamprey Acly"			
/note="Vector: pGEM-T Easy; lymphocyte mRNA ESTs from PCR			
subtracted cDNA libraries of immune stimulated larvae. All			
are single pass 5' or 3' sequences randomly cloned in			
pGEM-T Easy (Promega)."			
Location/Qualifiers			
1..1048			
/organism="Petromyzon marinus"			
/mol_type="mRNA"			
/db_xref="taxon:7757"			
/cell_type="lymphocyte"			
/dev_stage="immune stimulated larvae"			
/clone_lib="Sea lamprey Acly"			
/note="Vector: pGEM-T Easy; lymphocyte mRNA ESTs from PCR			
subtracted cDNA libraries of immune stimulated larvae. All			
are single pass 5' or 3' sequences randomly cloned in			
pGEM-T Easy (Promega)."			
Location/Qualifiers			
1..1048			
/organism="Petromyzon marinus"			
/mol_type="mRNA"			
/db_xref="taxon:7757"			
/cell_type="lymphocyte"			
/dev_stage="immune stimulated larvae"			
/clone_lib="Sea lamprey Acly"			
/note="Vector: pGEM-T Easy; lymphocyte mRNA ESTs from PCR			
subtracted cDNA libraries of immune stimulated larvae. All			
are single pass 5' or 3' sequences randomly cloned in			
pGEM-T Easy (Promega)."			
Location/Qualifiers			
1..1048			
/organism="Petromyzon marinus"			
/mol_type="mRNA"			
/db_xref="taxon:7757"			
/cell_type="lymphocyte"			
/dev_stage="immune stimulated larvae"			
/clone_lib="Sea lamprey Acly"			
/note="Vector: pGEM-T Easy; lymphocyte mRNA ESTs from PCR			
subtracted cDNA libraries of immune stimulated larvae. All			
are single pass 5' or 3' sequences randomly cloned in			
pGEM-T Easy (Promega)."			
Location/Qualifiers			
1..1048			
/organism="Petromyzon marinus"			
/mol_type="mRNA"			
/db_xref="taxon:7757"			
/cell_type="lymphocyte"			
/dev_stage="immune stimulated larvae"			
/clone_lib="Sea lamprey Acly"			
/note="Vector: pGEM-T Easy; lymphocyte mRNA ESTs from PCR			
subtracted cDNA libraries of immune stimulated larvae. All			
are single pass 5' or 3' sequences randomly cloned in			
pGEM-T Easy (Promega)."			
Location/Qualifiers			
1..1048			
/organism="Petromyzon marinus"			
/mol_type="mRNA"			
/db_xref="taxon:7757"			
/cell_type="lymphocyte"			
/dev_stage="immune stimulated larvae"			
/clone_lib="Sea lamprey Acly"			
/note="Vector: pGEM-T Easy; lymphocyte mRNA ESTs from PCR			
subtracted cDNA libraries of immune stimulated larvae. All			
are single pass 5' or 3' sequences randomly cloned in			
pGEM-T Easy (Promega)."			
Location/Qualifiers			
1..1048			
/organism="Petromyzon marinus"			
/mol_type="mRNA"			
/db_xref="taxon:7757"			
/cell_type="lymphocyte"			
/dev_stage="immune stimulated larvae"			
/clone_lib="Sea lamprey Acly"			
/note="Vector: pGEM-T Easy; lymphocyte mRNA ESTs from PCR			
subtracted cDNA libraries of immune stimulated larvae. All			
are single pass 5' or 3' sequences randomly cloned in			
pGEM-T Easy (Promega)."			
Location/Qualifiers			
1..1048			
/organism="Petromyzon marinus"			
/mol_type="mRNA"			
/db_xref="taxon:7757"			
/cell_type="lymphocyte"			
/dev_stage="immune stimulated larvae"			
/clone_lib="Sea lamprey Acly"			
/note="Vector: pGEM-T Easy; lymphocyte mRNA ESTs from PCR			
subtracted cDNA libraries of immune stimulated larvae. All			
are single pass 5' or 3' sequences randomly cloned in			
pGEM-T Easy (Promega)."			
Location/Qualifiers			
1..1048			
/organism="Petromyzon marinus"			
/mol_type="mRNA"			
/db_xref="taxon:7757"			
/cell_type="lymphocyte"			
/dev_stage="immune stimulated larvae"			
/clone_lib="Sea lamprey Acly"			
/note="Vector: pGEM-T Easy; lymphocyte mRNA ESTs from PCR			
subtracted cDNA libraries of immune stimulated larvae. All			
are single pass 5' or 3' sequences randomly cloned in			
pGEM-T Easy (Promega)."			
Location/Qualifiers			
1..1048			
/organism="Petromyzon marinus"			
/mol_type="mRNA"			
/db_xref="taxon:7757"			
/cell_type="lymphocyte"			
/dev_stage="immune stimulated larvae"			
/clone_lib="Sea lamprey Acly"			
/note="Vector: pGEM-T Easy; lymphocyte mRNA ESTs from PCR			
subtracted cDNA libraries of immune stimulated larvae. All			
are single pass 5' or 3' sequences randomly cloned in			
pGEM-T Easy (Promega)."			
Location/Qualifiers			
1..1048			
/organism="Petromyzon marinus"			
/mol_type="mRNA"			
/db_xref="taxon:7757"			
/cell_type="lymphocyte"			
/dev_stage="immune stimulated larvae"			
/clone_lib="Sea lamprey Acly"			
/note="Vector: pGEM-T Easy; lymphocyte mRNA ESTs from PCR			
subtracted cDNA libraries of immune stimulated larvae. All			
are single pass 5' or 3' sequences randomly cloned in			
pGEM-T Easy (Promega)."			
Location/Qualifiers			
1..1048			
/organism="Petromyzon marinus"			
/mol_type="mRNA"			
/db_xref="taxon:7757"			
/cell_type="lymphocyte"			
/dev_stage="immune stimulated larvae"			
/clone_lib="Sea lamprey Acly"			
/note="Vector: pGEM-T Easy; lymphocyte mRNA ESTs from PCR			
subtracted cDNA libraries of immune stimulated larvae. All			
are single pass 5' or 3' sequences randomly cloned in			
pGEM-T Easy (Promega)."			
Location/Qualifiers			
1..1048			
/organism="Petromyzon marinus"			
/mol_type="mRNA"			
/db_xref="taxon:7757"			
/cell_type="lymphocyte"			
/dev_stage="immune stimulated larvae"			
/clone_lib="Sea lamprey Acly"			
/note="Vector: pGEM-T Easy; lymphocyte mRNA ESTs from PCR			
subtracted cDNA libraries of immune stimulated larvae. All			
are single pass 5' or 3' sequences randomly cloned in			
pGEM-T Easy (Promega)."			
Location/Qualifiers			
1..1048			
/organism="Petromyzon marinus"			
/mol_type="mRNA"			
/db_xref="taxon:7757"			
/cell_type="lymphocyte"			
/dev_stage="immune stimulated larvae"			
/clone_lib="Sea lamprey Acly"			
/note="Vector: pGEM-T Easy; lymphocyte mRNA ESTs from PCR			
subtracted cDNA libraries of immune stimulated larvae. All			
are single pass 5' or 3' sequences randomly cloned in			
pGEM-T Easy (Promega)."			
Location/Qualifiers			
1..1048			
/organism="Petromyzon marinus"			
/mol_type="mRNA"			
/db_xref="taxon:7757"			
/cell_type="lymphocyte"			
/dev_stage="immune stimulated larvae"			
/clone_lib="Sea lamprey Acly"			
/note="Vector: pGEM-T Easy; lymphocyte mRNA ESTs from PCR			
subtracted cDNA libraries of immune stimulated larvae. All			
are single pass 5' or 3' sequences randomly cloned in			
pGEM-T Easy (Promega)."			
Location/Qualifiers			
1..1048			
/organism="Petromyzon marinus"			
/mol_type="mRNA"			
/db_xref="taxon:7757"			
/cell_type="lymphocyte"			
/dev_stage="immune stimulated larvae"			
/clone_lib="Sea lamprey Acly"			
/note="Vector: pGEM-T Easy; lymphocyte mRNA ESTs from PCR			
subtracted cDNA libraries of immune stimulated larvae. All			
are single pass 5' or 3' sequences randomly cloned in			
pGEM-T Easy (Promega)."			
Location/Qualifiers			
1..1048			
/organism="Petromyzon marinus"			
/mol_type="mRNA"			
/db_xref="taxon:7757"			
/cell_type="lymphocyte"			
/dev_stage="immune stimulated larvae"			
/clone_lib="Sea lamprey Acly"			
/note="Vector: pGEM-T Easy; lymphocyte mRNA ESTs from PCR			
subtracted cDNA libraries of immune stimulated larvae. All			
are single pass 5' or 3' sequences randomly cloned in			
pGEM-T Easy (Promega)."			
Location/Qualifiers			
1..1048			
/organism="Petromyzon marinus"			
/mol_type="mRNA"			
/db_xref="taxon:7757"			
/cell_type="lymphocyte"			
/dev_stage="immune stimulated larvae"			
/clone_lib="Sea lamprey Acly"			
/note="Vector: pGEM-T Easy; lymphocyte mRNA ESTs from PCR			
subtracted cDNA libraries of immune stimulated larvae. All			
are single pass 5' or 3' sequences randomly cloned in			
pGEM-T Easy (Promega)."			
Location/Qualifiers			
1..1048			
/organism="Petromyzon marinus"			
/mol_type="mRNA"			
/db_xref="taxon:7757"			
/cell_type="lymphocyte"			
/dev_stage="immune stimulated larvae"			
/clone_lib="Sea lamprey Acly"			
/note="Vector: pGEM-T Easy; lymphocyte mRNA ESTs from PCR			
subtracted cDNA libraries of immune stimulated larvae. All			
are single pass 5' or 3' sequences randomly cloned in			
pGEM-T Easy (Promega)."			
Location/Qualifiers			
1..1048			
/organism="Petromyzon marinus"			
/mol_type="mRNA"			
/db_xref="taxon:7757"			
/cell_type="lymphocyte"			
/dev_stage="immune stimulated larvae"			
/clone_lib="Sea lamprey Acly"			
/note="Vector: pGEM-T Easy; lymphocyte mRNA ESTs from PCR			
subtracted cDNA libraries of immune stimulated larvae. All			
are single pass 5' or 3' sequences randomly cloned in			
pGEM-T Easy (Promega)."			
Location/Qualifiers			
1..1048			
/organism="Petromyzon marinus"			
/mol_type="mRNA"			

```
QY 2260 GGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGTATGTAGGCGGTGCTACAGAGTT 2319
Db 420 GGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGTATGTAGGCGGTGCTACAGAGTT 479
QY 2320 CTTGAGAGTGGTGGCTTAACCTACCGCTACACTAGAGCAGAGTATTTGGTATCTCGCTCT 2379
Db 480 CTTGAGAGTGGTGGCTTAACCTACCGCTACACTAGAGCAGAGTATTTGGTATCTCGCTCT 539
QY 2380 GCTGAAGCCAGTTACCTTCGAAAAAGAGTTGTGTAGCTCTTGATCCGGCAAAACCAAC 2439
Db 540 GCTGAAGCCAGTTACCTTCGAAAAAGAGTTGTGTAGCTCTTGATCCGGCAAAACCAAC 599
QY 2440 CGCTGTAGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2499
Db 600 CGCTGTAGCG--TGTTTTTTTGTTCGCAAGCAGCAGATTACCGCGCAGAAAAAAGGATC 657
QY 2500 TCAAGAGAGATCCTTTTCATCTTTCTACGGGTCTGACGCTCAGTGGAGCAGAACTCAGG 2559
Db 658 TCAAGAGAGATCCTTTTCATCTTTCTACGGGTCTGACGCTCAGTGGAGCAGAACTCAGG 717
QY 2560 TTAAGGGATTTTGGTTCATGAGATTATCAAAAAAGGATCTTCACCTAGATCCTTTTAAATTA 2619
Db 718 TTAAGGGATTTTGGTTCATGAGATTATCAAAAAAGGATCTTCACCTAGATCCTTTTAAATTA 777
QY 2620 AAAATGAAGTTTAAATCAATCTAAAGTATATATGAGTAAACTTTGGTCTGACAGTTACCA 2679
Db 778 AAAATGAAGTTTAAATCAATCTAAAGTATATATGAGTAAACTTTGGTCTGACAGTTACCA 837
QY 2680 ATGCTTAATCAGTGAGGACCTATCTCAGGATCTCTATTTCTATTTCTATCTATCTATCTG 2739
Db 838 ATGCTTAATCAGTGAGGACCTATCTCAGGATCTCTATTTCTATTTCTATCTATCTATCTG 897
QY 2740 CTGACTCCCGCTCGTGTAGATACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGC 2799
Db 898 CTGACTCCCGCTCGTGTAGATACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGC 957
QY 2800 TGAATGATACCGCGAGACCCAGCTCAGCGGTCCAGATTTATACGCAATAAACCAAGCC 2859
Db 958 TGAATGATACCGCGAGACCCAGCTCAGCGGTCCAGATTTATACGCAATAAACCAAGCC 1017
QY 2860 AGCCGAAGGGCCGAGCGCAGAGTGTCT 2890
Db 1018 AGCCGAAGGGCCGAGCGCAGAGTGTCT 1048

RESULT 3
BM438846/c
LOCUS
DEFINITION
  Ictalurus punctatus (channel catfish)
ACCESSION
  BM438846
VERSION
  EST.
KEYWORDS
  Ictalurus punctatus (channel catfish)
SOURCE
  Ictalurus punctatus
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes;
  Ictaluridae; Ictalurus.
  1 (bases 1 to 1013)
AUTHORS
  Feng, J., Kucuktas, H., Kocabas, A., Li, P. and Liu, Z.
TITLE
  Transcriptome of channel catfish (Ictalurus punctatus): initial
  analysis of expressed sequence tags from the liver
JOURNAL
  Unpublished (2002)
COMMENT
  Contact: Liu ZJ
  The Fish Molecular Genetics and Biotechnology Laboratory,
  Department of Fisheries and Allied Aquacultures and Program of Cell
  and Molecular Biosciences
  Auburn University
  203 Swingle Hall, Auburn University, Auburn, AL 36849, USA
  Tel: 334 844 4054
  Fax: 334 844 9208
  Email: zliu@acesag.auburn.edu
```

```
Seq primer: M13 Reverse.
Location/Qualifiers
  1. .1013
  /organism="Ictalurus punctatus"
  /mol_type="mRNA"
  /db_xref="taxon:7998"
  /clone_lib="Liver cDNA library"
  /note="Organ: Liver; Vector: pSport1; Site_1: NotI;
  Site_2: Sall"

ORIGIN
Query Match      26.9%; Score 986.4; DB 4; Length 1013;
Best Local Similarity 99.5%; Pred. No. 1.1e-278;
Matches 1009; Conservative 1; Mismatches 2; Indels 2; Gaps 2;

QY 2477 ATTACCGCGCAGAAAAAAGGATCTCAAGAGAGTCTCTTGGATCTTTTCTACGGGTCTGAC 2536
Db 1013 ATTACCGCGCAGAAAAAAGGATCTCAAGAGAGTCTCTTGGATCTTTTCTACGGGTCTGAC 954
QY 2537 GCTCAGTGAACGAAAACTCAGCTTAAGGGATTTTGGTCAATGAGATTATCAAAAAGATC 2596
Db 953 GCTCAGTGAACGAAAACTCAGCTTAAGGGATTTTGGTCAATGAGATTATCAAAAAGATC 894
QY 2597 TTCACCTAGATCCTTTTAAATTAATAATGAAGTTTAAATCAATCTAAAGTATATATGAG 2656
Db 893 TTCACCTAGATCCTTTTAAATTAATAATGAAGTTTAAATCAATCTAAAGTATATATGAG 834
QY 2657 TAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGT 2716
Db 833 TAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGT 774
QY 2717 CTATTTTCGTTTCATTCATAGTTGCTGATCTCCCGTCTGTAGATAAATACGATACGGGAG 2776
Db 773 CTATTTTCGTTTCATTCATAGTTGCTGATCTCCCGTCTGTAGATAAATACGATACGGGAG 714
QY 2777 GGCTTACCATCTGGCCCCAGTCTGCAATGATACCGGAGACCCACGCTCACCAGTCTCA 2836
Db 713 GGCTTACCATCTGGCCCCAGTCTGCAATGATACCGGAGACCCACGCTCACCAGTCTCA 654
QY 2837 GATTTATCAGCAATAAACCCAGCAGCCGGAAGGGCCGAGCGAGAGTGTCTCGCAACT 2896
Db 653 GATTTATCAGCAATAAACCCAGCAGCCGGAAGGGCCGAGCGAGAGTGTCTCGCAACT 594
QY 2897 TTATCCGCTCCATCCAGTCTATTAATTTGTTCGGGAAGCTAGAGTAAGTAGTTCGCCA 2956
Db 593 TTATCCGCTCCATCCAGTCTATTAATTTGTTCGGGAAGCTAGAGTAAGTAGTTCGCCA 534
QY 2957 GTTAATAGTTTGGCAACGTTGTGTCATTTGCTACAGGCATCGTGTGTACCGCTCGTCG 3016
Db 533 GTTAATAGTTTGGCAACGTTGTGTCATTTGCTACAGGCATCGTGTGTACCGCTCGTCG 474
QY 3017 TTTGGTATGGCTTCATTTCAGCTCCGGTTCCTCAACGATCAAGGCGAGTTACA-TGATCCCC 3075
Db 473 TTTGGTATGGCTTCATTTCAGCTCCGGTTCCTCAACGATCAAGGCGAGTTACATTGATCCCC 414
QY 3076 CATGTTGTCAAAAAAGCGTTAGTCTTCCTCGGTCTCCGATCGTTGTTCAGAAAGTAGTT 3135
Db 413 CATGTTGTCAAAAAAGCGTTAGTCTTCCTCGGTCTCCGATCGTTGTTCAGAAAGTAGTT 355
QY 3136 GGCCGAGTGTATCACTCATGTTATGGCAGCAGTGCATTAATCTTACTCTCATGCC 3195
Db 354 GGCCGAGTGTATCACTCATGTTATGGCAGCAGTGCATTAATCTTACTCTCATGCC 295
QY 3196 ATCCGTAAGATGCTTTTCTGTGACTGCTGAGTACTCAACCAAGTCAATCTTGAGAATAGTG 3255
Db 294 ATCCGTAAGATGCTTTTCTGTGACTGCTGAGTACTCAACCAAGTCAATCTTGAGAATAGTG 235
QY 3256 TATGCGGACCGAGTGTCTTTGCGCGGTCAATACGGGATAATACCGCGCACATAG 3315
Db 234 TATGCGGACCGAGTGTCTTTGCGCGGTCAATACGGGATAATACCGCGCACATAG 175
QY 3316 CAGAATTTAAAGTGTCTCATCTATCTGAAAAAGTCTTCTGGGCGGAAAACTCTCAAGAT 3375
Db 3316 CAGAATTTAAAGTGTCTCATCTATCTGAAAAAGTCTTCTGGGCGGAAAACTCTCAAGAT 3375
```

Dn	174	CAGAACTTTAAAGTGCTCATCATTTGGAAAAAGTTCTTCGGGCGCAAAACTCTCAAGGAT	115
Qy	3376	CTTACCGCTGTTGAGATCCAGTTCGATGAACCACTCGTGCACCAACTGATCTTCAGC	3435
Dn	114	CTTACCGCTGTTGAGATCCAGTTCGATGAACCACTCGTGCACCAACTGATCTTCAGC	55
Qy	3436	ATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAACAGGAAGCAAAATGC	3489
Dn	54	ATCTTTAACTTTTACCAGCGTCTCTGGGTGAGCAAAACAGGAGCAAAATGC	1
RESULT 4			
AU081137			
LOCUS	AU081137	1067 bp mRNA linear EST 30-JUL-2002	
DEFINITION	AU081137	Oncorhynchus mykiss Kidney infected by infectious hematopoietic necrosis virus Oncorhynchus mykiss cDNA clone KI2, mRNA sequence.	
ACCESSION	AU081137		
VERSION	AU081137.1	GI:6431485	
KEYWORDS			
SOURCE		Oncorhynchus mykiss (rainbow trout)	
ORGANISM		Oncorhynchus mykiss	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.	
AUTHORS		Kono, T., Sakai, M. and Lapetra, S.E.	
TITLE		Expressed Sequence Tag Analysis of Kidney and Gill Tissues from Rainbow Trout (Oncorhynchus mykiss) Infected with Infectious Hematopoietic Necrosis Virus	
JOURNAL		Mar. Biotechnol. 2 (5), 493-498 (2001)	
COMMENT		Contact: Masahiro Sakai Faculty of Agriculture Miyazaki University 1-1 niishi gakukenbanadai, Miyazaki, Miyazaki 889-2192, Japan Email: m.sakai@cc.miyazaki-u.ac.jp.	
FEATURES	source	1..1067 Location/Qualifiers /organism="Oncorhynchus mykiss" /mol_type="mRNA" /db_xref="taxon:8022" /clone="KI2" /tissue_type="Kidney" /clone_lib="Oncorhynchus mykiss Kidney infected by infectious hematopoietic necrosis virus" /note="common name:rainbow trout ; infected by infectious hematopoietic necrosis virus"	
ORIGIN			
Query Match		26.1%; Score 955.2; DB 1; Length 1067;	
Best Local Similarity		99.0%; Pred. No. 1.8e-269;	
Matches 982; Conservative		0; Mismatches 8; Indels 2; Gaps 2;	
Qy	1494	CGCGTTGCGGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTTGTTATCGGTCACAA	1553
Dn	78	CGAGCTGCGGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTTGTTATCGGTCACAA	137
Qy	1554	TTCCACACACATACGACCGCAAGCATAAAGTGTAAGCCCTGGGTCCTTAATGAGTGA	1613
Dn	138	TTCCACACACATACGACCGCAAGCATAAAGTGTAAGCCCTGGGTCCTTAATGAGTGA	197
Qy	1614	GCTAACTCACATTAATTCGCTGCTCACTCCCGCTTTCCAGTCGGGAACCTGTGCT	1673
Dn	198	GCTAACTCACATTAATTCGCTGCTCACTCCCGCTTTCCAGTCGGGAACCTGTGCT	257
Qy	1674	GCCAGCTGATTAATGAATCGGCCAACCGCGGGGAGAGCGGTTTGGGTATTTGGCGCT	1733
Dn	258	GCCAGCTGATTAATGAATCGGCCAACCGCGGGGAGAGCGGTTTGGGTATTTGGCGCT	317
Qy	1734	CTTCCGCTTCCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT	1793
Dn	318	CTTCCGCTTCCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT	377
Qy	1794	CAGCTCACTCAAAGCGGTAAATACGGTTTATCCACAGAATCAGGGGATACGAGGAAGA	1853
Dn	378	CAGCTCACTCAAAGCGGTAAATACGGTTTATCCACAGAATCAGGGGATACGAGGAAGA	437
Qy	1854	ACATGTGAGCAAAAGGCGCAGCAAAAGGCGCAGAAACCGTAAAGGCGGTTGTTGGCGT	1913
Dn	438	ACATGTGAGCAAAAGGCGCAGCAAAAGGCGCAGAAACCGTAAAGGCGGTTGTTGGCGT	497
Qy	1914	TTTTTCATAGGCTCGCCCGCTGACGAGCATCAAAAATCGAGCTCAAGTCAGAGGT	1973
Dn	498	VVTTTCATAGGCTCGCCCGCTGACGAGCATCAAAAATCGAGCTCAAGTCAGAGGT	557
Qy	1974	GCGAAACCCGACGAGCACTATAAAGATACAGGCGCTTTCCCGCTCGAAAGCTCCCTCGTGC	2033
Dn	558	GCGAAACCCGACGAGCACTATAAAGATACAGGCGCTTTCCCGCTCGAAAGCTCCCTCGTGC	617
Qy	2034	GCTCTCTGTTCCGACCCCTGCGGTTACGGATACCTGTCCGCCCTTTCTCCCTCGGGAA	2093
Dn	618	GCTCTCTGTTCCGACCCCTGCGGTTACGGATACCTGTCCGCCCTTTCTCCCTCGGGAA	677
Qy	2094	GCGTGGCGCTTCTCATAGCTCAGCTGATGATATCTCAGTTTCGTTAGGTGCTTCGCT	2153
Dn	678	GCGTGGCGCTTCTCATAGCTCAGCTGATGATATCTCAGTTTCGTTAGGTGCTTCGCT	737
Qy	2154	CAAAGCTGGGCTGTGTGCAAGAAACCCCGCTTTCAGCCGACCGCTTCGCTTATCCGATA	2213
Dn	738	CAAAGCTGGGCTGTGTGCAAGAAACCCCGCTTTCAGCCGACCGCTTCGCTTATCCGATA	797
Qy	2214	ACTATCGTCTTGAATCCAAACCGGTAAGACAGCACTATCGCCACTGGCAGCAGCACTG	2273
Dn	798	ACTATCGTCTTGAATCCAAACCGGTAAGACAGCACTATCGCCACTGGCAGCAGCACTG	857
Qy	2274	GTAACAGGATTAAGCAGAGGATATGTAGCGGTGTACAGAGTTCTTGAAGTGTGTGC	2333
Dn	858	GTAACAGGATTAAGCAGAGGATATGTAGCGGTGTACAGAGTTCTTGAAGTGTGTGC	917
Qy	2334	CTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTCGGCTCTGCTGAAGCCAGTTA	2393
Dn	918	CTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTCGGCTCTGCTGAAGCCAGTTA	977
Qy	2394	CCTTCGAAAAAGAGTTGGTAGCTCTTGTATCGGCAAAACAAACACCGGTGTAGCGGTG	2453
Dn	978	CCTTCGAAAAAGAGTTGGTAGCTCTTGTATCGGCAAAACAAACCGGTGTAGCGGTG	1036
Qy	2454	GTCTTTTGTGTAAGCAGCAGATTTACGCGC	2485
Dn	1037	TTTTTTTGTGTC-AGCAGCAGATTACGCGC	1067
RESULT 5			
AJ281480			
LOCUS	AJ281480	1004 bp mRNA linear EST 30-JUN-2000	
DEFINITION	AJ281480	4A3A-P4G8-F Anopheles gambiae immune competent 4A3A Anopheles gambiae cDNA clone 4A3A-P4G8, mRNA sequence.	
ACCESSION	AJ281480		
VERSION	AJ281480.1	GI:6929360	
KEYWORDS		EST.	
SOURCE		Anopheles gambiae (African malaria mosquito)	
ORGANISM		Anopheles gambiae	
REFERENCE		Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.	
AUTHORS		1 (bases 1 to 1004) Dimopoulos, G., Casavant, T.L., Chang, S., Scheetz, T., Roberts, C., Donohue, M., Schultz, J., Benes, V., Bork, P., Ansorge, W., Soares, M.B. and Kafatos, F.C.	
TITLE		Anopheles gambiae pilot gene discovery project: identification of mosquito innate immunity genes from expressed sequence tags generated from immune-competent cell lines	
JOURNAL		Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)	
MEDLINE		20300950	
PUBMED		10841561	
COMMENT		Contact: Dimopoulos G	

Db	661	ACGCTCAGTGGGAACGAAACATCACGTTAAGGGAATTTTGGTCATGAGATTATCAAAAGGA	720
Qy	2595	TCTTCACCTAGATCCTTTTAAATTTAAAAATGAAGTTTTTAAATCAATCTTAAAGTATATATG	2654
Db	721	TCTTCACCTAGATCCTTTTAAATTTAAAAATGAAGTTTTTAAATCAATCTTAAAGTATATATG	780
Qy	2655	AGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAAGCCACCTATCTCAGCGATCT	2714
Db	781	AGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAAGCCACCTATCTCAGCGATCT	839
Qy	2715	GTCTATTTCGTTCCATCCATAGTTGCTGACTCCCTCCGTCGTTAGTAGTAACCTACCATACGGG	2774
Db	840	GTCTATTTCGTTCCATCCATAGTTGCTGACTCCCTCCGTCGTTAGTAGTACTACCATACGGG	898
Qy	2775	AGGGCTTACCATCTGCGCCCGAGTCTGCTCAATGATACCGGAGACCCACGCTACCCGGCTC	2834
Db	899	AGGGCTTACCATCTGCGCCCGAGTCTGCTCAATGATACCGGAGACCCACGCTACCCGGCTC	955
Qy	2835	CAGATTATCAGCAATAAACCCAGCCGGAAGGGCCGAGCCAGAGTG	2885
Db	956	CAGATTATCAGCAATAAACCCAGCCGGAAGGGCCGAGCCAGAGTG	1004
RESULT 6			
LOCUS	AU081124		
DEFINITION	AU081124	Oncorhynchus mykiss Kidney infected by infectious hematopoietic necrosis virus	EST 30-JUL-2002
ACCESSION	AU081124		
VERSION	AU081124.1	GI:6431472	
KEYWORDS	EST.		
SOURCE	Oncorhynchus mykiss (rainbow trout)		
ORGANISM	Oncorhynchus mykiss		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.		
AUTHORS	1 (bases 1 to 1089)		
TITLE	Kono,T., Sakai,M. and LaPatra,S.E.		
JOURNAL	Expressed Sequence Tag Analysis of Kidney and Gill Tissues from Rainbow Trout (Oncorhynchus mykiss) Infected with Infectious Hematopoietic Necrosis Virus		
COMMENT	Mar. Biotechnol. 2 (5), 493-498 (2001)		
FEATURES	Contact: Masahiro Sakai		
source	Faculty of Agriculture		
	Miyazaki University		
	1-1 nishi sakuenkibanadai, Miyazaki, Miyazaki 889-2192, Japan		
	Email: m.sakae@cc.miyazaki-u.ac.jp.		
	Location/Qualifiers		
	1..1089		
	/organism="Oncorhynchus mykiss"		
	/mol_type="mRNA"		
	/db_xref="taxon:8022"		
	/clone="KG-12"		
	/tissue_type="kidney"		
	/clone_lib="Oncorhynchus mykiss Kidney infected by infectious hematopoietic necrosis virus"		
	/notes="common name:rainbow trout ; infected by infectious hematopoietic necrosis virus"		
ORIGIN			
Query Match	25.2%;	Score 923.8;	DB 1; Length 1089;
Best Local Similarity	97.2%;	Pred. No. 3.3e-260;	
Matches	983; Conservative	0; Mismatches	22; Indels 6; Gaps 4;
Qy	1494	CGCGTGGCGTAAATCATGGTCATAGCTGTTTCTCTGTGTGAATTTGTTATCCGCTCACAA	1553
Db	79	CGAGCTTGGCGTAAATCATGGTCATAGCTGTTTCTCTGTGTGAATTTGTTATCCGCTCACAA	138
Qy	1554	TTCCACACACATACAGCCGGAAGCATAAAGTGTAAGCCCTGGGCTGCTTAATGAGTGA	1613
Db	139	TTCCACACACATACAGCCGGAAGCATAAAGTGTAAGCCCTGGGCTGCTTAATGAGTGA	198

[illegible]

Xenopodinae; Xenopus; Silurana.
1 (bases 1 to 1049)
Kremitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
Mardis,E. and Wilson.R.
A physical map of the xenopus tropicalis genome
Unpublished (2003)
Contact: Richard K Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submission@watson.wustl.edu
Insert Length: 175000 Std Error: 0.00
Seq primer: RML TAGGACTCACTATAGGAGA
Class: BAC ends
High quality sequence start: 43
High quality sequence stop: 888.
Location/Qualifiers
1..1049
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/strain="Nigerian frog"
/db_xref="taxon:8364"
/clone="CH216-8A14"
/sex="male"
/cell_line="Stock 248 F7A2, inbred N7"
/clone_lib="CH216"
/notes="vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis
BAC library"

ORIGIN

Query Match	24.7%	Score	902.8	DB	9	Length	1049		
Best Local Similarity	98.0%	Pred. No.	5.2e-254						
Matches	935	Conservative	0	Mismatches	17	Indels	2	Gaps	2
Qy	1486	CCGCGCGGCGCGCTTGGCGTAAATCATGGTCATAGCTGTTTCCGTGTGGAATTTGTTATCC	1545						
Db	95	CAGCATGCAAGCTTGGCGTAAATCATGGTCATAGCTGTTTCCGTGTGGAATTTGTTATCC	154						
Qy	1546	GCTCACAATTTCCACACAACATACGAGCCGGAAGATAAAGTGTAAGCCTTGGGGTGCCCTA	1605						
Db	155	GCTCACAATTTCCACACAACATACGAGCCGGAAGATAAAGTGTAAGCCTTGGGGTGCCCTA	214						
Qy	1606	ATGAGTGAGTAAGTACATTAATTTGGGTGCGCTCACTGCCCGCTTTCAGTTCGGGAAA	1665						
Db	215	ATGAGTGAGTAAGTACATTAATTTGGGTGCGCTCACTGCCCGCTTTCAGTTCGGGAAA	274						
Qy	1666	CCTGTGCTGCCAGCTGCATTAATGAATTCGGCCAAACGCGGGGAGAGCGCGTTTGCGTAT	1725						
Db	275	CCTGTGCTGCCAGCTGCATTAATGAATTCGGCCAAACGCGGGGAGAGCGCGTTTGCGTAT	334						
Qy	1726	TGGCGCGCTTTTCGCGTTCCTCGCTCACTGACTCGCTCGCGTTCGGTTCGGTTCGGCGG	1785						
Db	335	TGGCGCGCTTTTCGCGTTCCTCGCTCACTGACTCGCTCGCGTTCGGTTCGGTTCGGCGG	394						
Qy	1786	AGCGGTATCAGCTCACTCAAAAGCGGTAAATACGGTTTATCCACAGAATCAGGGGATTAACGC	1845						
Db	395	AGCGGTATCAGCTCACTCAAAAGCGGTAAATACGGTTTATCCACAGAATCAGGGGATTAACGC	454						
Qy	1846	AGGAAAGAAACATGTGAGCAAAAGCCAGCAAAAGCCAGGAAACCGTAAAAAGCCCGGTT	1905						
Db	455	AGGAAAGAAACATGTGAGCAAAAGCCAGCAAAAGCCAGGAAACCGTAAAAAGCCCGGTT	514						
Qy	1906	GCTGGCGTTTTTCCATAGGCTCGCGCCCTCTGACGAGCATCAAAAAATCGACGCTCAAG	1965						
Db	515	GCTGGCGTTTTTCCATAGGCTCGCGCCCTCTGACGAGCATCAAAAAATCGACGCTCAAG	574						
Qy	1966	TCAGAGTGGCGAAAACCCGACAGGACTATAAAGATACGAGGCGTTTTCCCTCGGAAGCTC	2025						
Db	575	TCAGAGTGGCGAAAACCCGACAGGACTATAAAGATACGAGGCGTTTTCCCTCGGAAGCTC	634						
Qy	2026	CCTCGTCEGCTCTCCTGTTCCGACCTTCGCGCTTACCGGATACCTGTCGCGCTTTCTCC	2085						
Db	635	CCTCGTCEGCTCTCCTGTTCCGACCTTCGCGCTTACCGGATACCTGTCGCGCTTTCTCC	694						


```
/organism="Pongo pygmaeus"
/mol_type="mRNA"
/db_xref="taxon:9600"
/clone="DKF2p469G045"
/tissue_type="kidney"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="469 (synonym: pkid1)"
/notes="Vector: pSport1_Sfi; Site_1: SfiI; Site_2: SfiI"

ORIGIN
Query Match      23.9%; Score 876.4; DB 7; Length 902;
Best Local Similarity 99.1%; Pred. No. 3e-246; 1; Indels 2; Gaps 2;
Matches 899; Conservative 0; Mismatches 1;

QY 2758 GATAACTACGATACGGAGGGCTTACCATCTGCGCCAGTGTCTGCAATGATACCGGAGA 2817
DB 901 GATAACTACGATACGGAGGGCTTACCATCTGCGCCAGTGTCTGCAATGATACCGGAGA 842
QY 2818 CCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGCCGAGCG 2877
DB 841 CCACGCTCACCGGCTCCAGATTTTTCAGCAATAAACCAGCCAGCCGGAAGGCCGAGCG 782
QY 2878 CAGAAGTGTCTGCAACTTATCCGCTCCATCCAGTCTATTAAATTGTCGCGGAAGC 2937
DB 781 CAGAAGTGTCTGCAACTTATCCGCTCCATCCAGTCTATTAAATTGTCGCGGAAGC 722
QY 2938 TAGAGTAAGTAGTTCGCCAGTTAATAGTTTGGCAACGGTTGTGCCATTTGCTACAGGCAT 2997
DB 721 TAGAGTAAGTAGTTCGCCAGTTAATAGTTTGGC-AACGTTGTGCCATTTGCTACAGGCAT 663
QY 2998 CTTGGTGTCAAGCTCGTCTGTTGGTATGCTTCATTCAGCTCCGGTTCACACGATCAAG 3057
DB 662 CTTGGTGTCAAGCTCGTCTGTTGGTATGCTTCATTCAGCTCCGGTTCACACGATCAAG 603
QY 3058 GCGAGTTACATGATCCCGCATGTTGTGCAAAAAGCGTTAGCTCTTGGTCTCTCGAT 3117
DB 602 GCGAGTTACATGATCCCGCATGTTGTGCAAAAAGCGTTAGCT-CTTGGTCTCTCGAT 544
QY 3118 CGTTGTGAGAAGTAAGTTGGCGCAGTGTATCACTCATGTTTATGGCAGCACTGCATAA 3177
DB 543 CGTTGTGAGAAGTAAGTTGGCGCAGTGTATCACTCATGTTTATGGCAGCACTGCATAA 484
QY 3178 TTCTCTTACTGTCAAGCATTCGTAAGATGCTTTTCTGTGACGTGTGAGTACTCAACCAA 3237
DB 483 TTCTCTTACTGTCAAGCATTCGTAAGATGCTTTTCTGTGACGTGTGAGTACTCAACCAA 424
QY 3238 GTCAATTCAGATAGTGTATGGCGGACCGAGTGTCTCTTGGCGGCGTCAATACGGA 3297
DB 423 GTCAATTCAGATAGTGTATGGCGGACCGAGTGTCTCTTGGCGGCGTCAATACGGA 364
QY 3298 TAATACCGCGCACATAGCAGAACTTTAAAGTGCTCATCATTTGGAAGAAAGTCTTCTCGG 3357
DB 363 TAATACCGCGCACATAGCAGAACTTTAAAGTGCTCATCATTTGGAAGAAAGTCTTCTCGG 304
QY 3358 GCGAAACTCTCAAGGATCTTACCGTGTGTAGATCCAGTTCGATGTAAACCACTCGTC 3417
DB 303 GCGAAACTCTCAAGGATCTTACCGTGTGTAGATCCAGTTCGATGTAAACCACTCGTC 244
QY 3418 ACCAACTGATCTTACGATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAACAGG 3477
DB 243 ACCAACTGATCTTACGATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAACAGG 184
QY 3478 AAGGCAAAATGCCGCAAAAGGAATAAGGGGACACGAAATGTTGAATCTCATACT 3537
DB 183 AAGGCAAAATGCCGCAAAAGGAATAAGGGGACACGAAATGTTGAATCTCATACT 124
QY 3538 CTTCTCTTTTCAATATTATTGAAGCATTTATCAGGGTTATTTGCTCATGAGCGGATACAT 3597
DB 123 CTTCTCTTTTCAATATTATTGAAGCATTTATCAGGGTTATTTGCTCATGAGCGGATACAT 64
QY 3598 ATTGAATGTATTAGAAAATAAACAATAGGGGTTCCGCGACATTTCCCGGAAAGT 3657
DB 64 ATTGAATGTATTAGAAAATAAACAATAGGGGTTCCGCGACATTTCCCGGAAAGT 3657
```

```
Db 63 ATTGAATGTATTAGAAAATAAACAATAGGGGTTCCGCGACATTTCCCGGAAAGT 4
QY 3658 GC 3659
DB 3 GC 2

RESULT 11
BG838279/c
LOCUS
DEFINITION
  Gc01_10e07_R Gc01_AAFc ECORC cold stressed Glycine clandestina
  Glycine clandestina cDNA clone Gc01_10e07, mRNA sequence.
ACCESSION
  BG838279
VERSION
  BG838279.1 GI:14204601
KEYWORDS
  EST.
SOURCE
  Glycine clandestina
ORGANISM
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
  Glycine.
REFERENCE
  1 (bases 1 to 935)
AUTHORS
  Singh, J.A., Farah, S., Chapados, J., Couroux, P., De Moors, A.,
  Harris, L.J., Hattori, J.I., Ouellet, T., Robert, L.S., Spratt, D. and
  Tinker, N.A.
TITLE
  Expressed Sequence Tags from Cold-Stressed Glycine clandestina
  Seedlings
JOURNAL
  Unpublished (2001)
COMMENT
  Contact: Singh, J.A.
  Eastern Cereal and Oilseed Research Centre
  Agriculture and Agri-food Canada
  KW Neatby Bldg., Central Experimental Farm, Ottawa, Ontario, K1A
  0C6, Canada
  Tel: (613) 759-1662
  Fax: (613) 759-1701
  Email: singhja@agr.gc.ca.
  Location/Qualifiers
    1..935
    /organism="Glycine clandestina"
    /mol_type="mRNA"
    /culture="1035"
    /db_xref="taxon:45687"
    /clone="Gc01_10e07"
    /tissue_type="Leaves, stem"
    /clone_lib="Gc01_AAFc_ECORC_cold_stressed_Glycine_clandest
    ina"
    /note="Vector: Bluescript SK+/XhoI-EcoRI; Site 1: EcoRI;
    Site 2: XhoI; Plants incubated at 2 degrees under 12 hours
    of light/day. Harvested after only 2-3 days of cold
    treatment. cDNA was prepared with the Uni-Zap cDNA kit
    from Stratagene. Eco RI adapters were linked followed by
    digest with Xho I/Eco RI and ligated to pBluescript."

FEATURES
  source
  Query Match      23.8%; Score 870.2; DB 4; Length 935;
  Best Local Similarity 98.1%; Pred. No. 2.1e-244;
  Matches 874; Conservative 14; Mismatches 2; Indels 1; Gaps 1;

QY 2773 GGAGGGCTTACCATCTGGCCAGTCTGCAATGATACCGAGAGCCAGCTCACCGG 2832
DB 928 GGAGGGCTTACCATCTGGCCAGTCTGCAATGATACCGAGAGCCAGCTCACCGG 869
QY 2833 TCCAGATTTTATCAGCAATAAACCCAGCCAGCGGAGGCGGAGAGTGGTCTGC 2892
DB 868 YCCAGATTTTATCAGCAATAAACCCAGCCAGCGGAGGCGGAGAGTGGTCTGC 809
QY 2893 AACTTTATCCGCTCCATCCAGTCTATTAAATTGTTCCCGGAGCTAGAGTAGTTC 2952
DB 808 AACTTTATCCGCTCCATCCAGTCTATTAAATTGTTCCCGGAGCTAGAGTAGTTC 749
QY 2953 GCGAGTTATAGTTTGGCAAGCTTGTTCCTTCCATTTGCTACAGGCATCGTGGTGCAGCTC 3012
DB 748 GCGAGTTATAGTTTGGCAAGCTTGTTCCTTCCATTTGCTACAGGCATCGTGGTGCAGCTC 689
```

```
QY 3013 GTCTGTTTGGTATGGCTTCATTGAGTCCCGTTCCCAACGATCAAGCGAGTTACATGATC 3072
D 688 GTCTGTTTGGTATGGCTTCATTGAGTCCCGTTCCCAACGATCARGCGAGTTACATGATC 629
QY 3073 CCCATGTTGTGCAAAAAGCGGTAGTCTTCGCTCCGATCCGATGTTGTGAGAGTAA 3132
D 628 CCCATGTTGTGCAAAAAGCGGTAGTCTTCGCTCCGATCCGATGTTGTGAGAGTAA 569
QY 3133 GTT-GGCCGCGAGTGTATCACTCATGTTTATGGGAGCACTGATATTTCTTACTGTCA 3191
D 568 GTTGGGCGCAGTGTATCACTCATGTTTATGGGAGCACTGATATTTCTTACTGTCA 509
QY 3192 TGCATCCGTAAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTCAGAA 3251
D 508 TGCATCCGTAAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTCAGAA 449
QY 3252 AGTGTATCGGCGACCGAGTTGCTCTTTCGCCGCGTCAATACGGGATATACCGGCCAC 3311
D 448 AGTGTATCGGCGACCGAGTTGCTCTTTCGCCGCGTCAATACGGGATATACCGGCCAC 389
QY 3312 ATAGCAGAACTTTAAAGTGCTCATCATTTGGAAGAGTTCTTCGGGGGAAACTCTCA 3371
D 388 ATAGCAGAACTTTAAAGTGCTCATCATTTGGAAGAGTTCTTCGGGGGAAACTCTCA 329
QY 3372 GGATCTTACCGCTGTGAGATCCAGTTCGATGTAACCCACTCGTGACCCCACTGATCT 3431
D 328 GGATCTTACCGCTGTGAGATCCAGTTCGATGTAACCCACTCGTGACCCCACTGATCT 269
QY 3432 CAGCATCTTTTACTTTACAGCGTTTCTGGGTGAGCAAAAACAGAGAGGCAAAATGCCG 3491
D 268 CAGCATCTTTTACTTTACAGCGTTTCTGGGTGAGCAAAAACAGAGAGGCAAAATGCCG 209
QY 3492 CAAAAAGGATATAGGCGACACGGATGTGATATCTACTCATCTCTCTTTTCAAT 3551
D 208 CAAAAAGGATATAGGCGACACGGATGTGATATCTACTCATCTCTCTTTTCAAT 149
QY 3552 ATTATTGAAGCATTTATCAGGGTATTGTCTCATGAGCGGATACATATTTGAATGTATT 3611
D 148 ATTATTGAAGCATTTATCAGGGTATTGTCTCATGAGCGGATACATATTTGAATGTATT 89
QY 3612 AGAAAAATAAACAATAGGGTTCCGCGACATTTCCCGGAAAGTGCCAC 3662
D 88 AGAAAAATAAACAATAGGGTTCCGCGACATTTCCCGGAAAGTGCCAC 38

RESULT 12
CF269652
LOCUS
DEFINITION
Fvylcolid844 Fragilariopsis cylindrus SMART cDNA library (Clontech)
Fragilariopsis cylindrus cDNA clone Antarctic 5', mRNA sequence.
ACCESSION
CF269652
VERSION
CF269652.1 GI:33631539
SOURCE
Fragilariopsis cylindrus
ORGANISM
Fragilariopsis cylindrus
Eukaryota; stramenopiles; Bacillariophyta; Bacillariophyceae;
Bacillariophycidae; Bacillariales; Bacillariaceae; Fragilariopsis.
1 (bases 1 to 1073)
Mock,T. and Valentin,K.
EST analysis of freezing tolerance in the Antarctic diatom
Fragilariopsis cylindrus: Detection of numerous cold adaptation
related genes and gene transfer events
Unpublished (2003)
Contact: Mock T
Biological Oceanography
Alfred-Wegener-Institute for Polar and Marine Research
Am Handelshafen 12, D-27570 Bremerhaven, Germany
Tel.: +49 471 4831 1893
Fax: +49 471 4831 1425
Email: tmock@awi-bremerhaven.de
sequence with unknown function
PCR PRIMERS
```

FORWARD: 5' lambdatriplex2
BACKWARD: 3' lambdatriplex2
Seq primer: ctcggaagcgcgccattgtgtgtgt.

FEATURES

source

1..1073
/organism="Fragilariopsis cylindrus"
/mol_type="mRNA"
/strain="Antarctic"
/db_xref="taxon:186039"
/clone="Antarctic"
/clone_lib="Fragilariopsis cylindrus SMART cDNA library (Clontech)"
/note="Vector: pTriplex2; total polyA was used for first-strand synthesis with SMART IV oligos and CDS III/3'PCR primer. Double strand cDNA synthesis was done by LD PCR using the following program: 95oC for 5 min denaturation and subsequent 20 cycles at 95oC (2min) and 68oC (6min). After SfiI digestion the cDNA was fractionated with CHROMA Spin-400 columns. These cDNAs were ligated overnight into pTriplex2 vectors."

ORIGIN

Query Match 23.6%; Score 865; DB 7; Length 1073;
Best Local Similarity 96.2%; Pred. No. 7.3e-243;
Matches 927; Conservative 0; Mismatches 30; Indels 7; Gaps 4;

QY 1631 GGGTTGCCCTCAGTCGCGCTTTCCAGTCGGGAAACCTGTCGTCAGCTGCATTAAATGA 1690
D 1 GCTTCTGCTACTGCCCGCTTTCAGTCGGGAAACCTGTCGTCAGCTGCATTAAATGA 60

QY 1691 ATGCGCCAAACGCGGGGAGAGCGGTTTGGGTATTGGCGCTTTCCGCTTCTTCGCTC 1750
D 61 ATGCGCCAAACGCGGGGAGAGCGGTTTGGGTATTGGCGCTTTCCGCTTCTTCGCTC 120

QY 1751 ACTGACTCGTCGCTCGGTGTTCCGCTGCGGAGCGGTATCAGCTCACTCAAGGCG 1810
D 121 ACTGACTCGTCGCTCGGTGTTCCGCTGCGGAGCGGTATCAGCTCACTCAAGGCG 180

QY 1811 GTAATACGGTTATCCACAGAAATCAGGGATTAACGAGGAAAGAAATGTGACAAAAGC 1870
D 181 GTAATACGGTTATCCACAGAAATCAGGGATTAACGAGGAAAGAAATGTGACAAAAGC 240

QY 1871 CAGCAAAAAGGCGAGAACCGTAAAAAAGCGGTTGTGGCGTTTTTTCATAGGCTCGCG 1930
D 241 CAGCAAAAAGGCGAGAACCGTAAAAAAGCGGTTGTGGCGTTTTTTCATAGGCTCGCG 300

QY 1931 CCCCTGACGAGCATCAAAAATCGAGCTCAAGTCAGAGTGGCGAAACCCGACAGA 1990
D 301 CCCCTGACGAGCATCAAAAATCGAGCTCAAGTCAGAGTGGCGAAACCCGACAGA 360

QY 1991 CTATAAGATACCGAGCGTTTCCCGCTCGAAGCTCCCTCGCTGCGCTCTCTCTTCCGACC 2050
D 361 CTATAAGATACCGAGCGTTTCCCGCTCGAAGCTCCCTCGCTGCGCTCTCTCTTCCGACC 420

QY 2051 CTGCGCTTACCGGATACCTGTGCGCCCTTCTCCCTTCGGGAAGCGTGGCGTTTCTCAT 2110
D 421 CTGCGCTTACCGGATACCTGTGCGCCCTTCTCCCTTCGGGAAGCGTGGCGTTTCTCAT 480

QY 2111 AGCTACGCTGTAGTATCTCAGTTCCGTTGATGTCGTTCCGTCGAAGTGGGCTGTGTG 2170
D 481 AGCTACGCTGTAGTATCTCAGTTCCGTTGATGTCGTTCCGTCGAAGTGGGCTGTGTG 540

QY 2171 CACGAAACCCCGTTTCAGCCGACCGCTTATCCGGTAACTATCTGCTTTCAGTCC 2230
D 541 CACGAAACCCCGTTTCAGCCGACCGCTTATCCGGTAACTATCTGCTTTCAGTCC 600

QY 2231 AACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGA 2290
D 601 AACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGA 660

QY 2291 GCGAGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCCT--AATACGGCTACA 2348
D 661 GCGAGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCNTAACTACNGGCTACA 720


```
QY 2349 CTAGAGGACA-GTATTGTGTATCT-GCGCTCTGCTGAAGCC---AGTTACCTTCGGAAA 2403
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 721 CTAGAGGAAACANGATATTGTGTATCTNGCGCTCTGCTNGAAGCCAGNTTACCNNTTCGAAA 780
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2404 AAGAGTTGGTAGCTCTTTGATCCGGCAACAAACACCGCTGGTAGCGGTGTTTTTGT 2463
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 781 AAGAGTTGGTAGCTCTTTGATCCGGCAACAAACACCGCTGGTAGCGGTGTTTTTGT 840
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2464 TTGCAAGCAGCAGATTACGGCGAGAAAAAAGGATCTCAAGAGATCCTTTGATCTTTTC 2523
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 841 TTGCAAGCAGCAGATTACGGCGAGAAAAAAGGATCTCAAGAGATCCTTTGATCTTTTC 900
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2524 TACGGGGTCTGACGCTCAGTGAACGAAACTCAGCTTAAGGATTTTGGTCAATGAGATT 2583
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 901 TACGGGGTCTGACGCTCAGTGAACGAAACTCAGCTTAAGGATTTTGGGATGAGACA 960
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2584 ATCA 2587
Db 961 ATAA 964

RESULT 13
LOCUS CL076016
DEFINITION CH216-138F20_RM1.1 CH216 Xenopus tropicalis genomic clone
ACCESSION CH216-138F20, genomic survey sequence.
VERSION CL076016
KEYWORDS CL076016.1 GI:40531929
SOURCE GSS.
ORGANISM Xenopus tropicalis (western clawed frog)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus; Silurana.
REFERENCE 1 (bases 1 to 885)
AUTHORS Kremitzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,
Mardis, E. and Wilson, R.
TITLE A physical map of the xenopus tropicalis genome
JOURNAL Unpublished (2003)
COMMENT Contact: Richard K Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Insert Length: 175000 Std Error: 0.00
Seq primer: RM1 TAGCACTCACTATAGGAGA
Class: BAC ends
High quality sequence start: 11
High quality sequence stop: 810.
FEATURES
source
1..885
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/strain="Nigerian frog"
/db_xref="taxon:8364"
/clone="CH216-138F20"
/sex="male"
/cell_line="Stock 248 F7A2, inbred N7"
/clone_lib="CH216"
/notes="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis
BAC library"

Query Match 23.4%; Score 857.6; DB 9; Length 885;
Best Local Similarity 99.3%; Pred. No. 1e-240; 5; Indels 1; Gaps 1;
Matches 871; Conservative 0; Mismatches 5;

QY 1535 AATTGTATCCGCTCAATTCACACATACGACCGGAGCATATAAGTGTAAAGCC 1594
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5 AATTGTATCCGCTCAATTCACACATACGACCGGAGCATATAAGTGTAAAGCC 64
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1595 TGGGGTGCTTATGAGTGAGCTCACTCAATTAATTGGTGGCTCACTGCCCGCTTC 1654
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
Db 65 TGGGGTGCTTATGAGTGAGCTCACTCAATTAATTGGTGGCTCACTGCCCGCTTC 124
QY 1655 CAGTCGGGAAAACCTGTCGTCGCCAGCTGCATTAATGAATCGGCCAAACGCGCGGAGAGGC 1714
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 125 CAGTCGGGAAAACCTGTCGTCGCCAGCTGCATTAATGAATCGGCCAAACGCGCGGAGAGGC 184
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1715 GGTTCGCTATTGGGGCGCTCTTCGCGCTTCTCGCTCACTGACTCGCTCGGCTCGGTCGTT 1774
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 185 GGTTCGCTATTGGGGCGCTCTTCGCGCTTCTCGCTCACTGACTCGCTCGGCTCGGTCGTT 244
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1775 CGGCTCGCGGAGCGGTATCAGCTCACTCAAAGCGGTAAATACGGTTATCCACAGAATCA 1834
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 245 CGGCTCGCGGAGCGGTATCAGCTCACTCAAAGCGGTAAATACGGTTATCCACAGAATCA 304
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1835 GGGGATAACCGCAGGAAAGAACATGTGAGCAAAAGGCGCAGCAAAAGCCAGGAAACCGTAAA 1894
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 305 GGGGATAACCGCAGGAAAGAACATGTGAGCAAAAGGCGCAGCAAAAGCCAGGAAACCGTAAA 364
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1895 AAGGCCGCGTTGCTGGCGCTTTTTCATAGGCTCCGCCCGCTGACGAGCATCAAAAT 1954
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 365 AAGGCCGCGTTGCTGGCGCTTTTTCATAGGCTCCGCCCGCTGACGAGCATCAAAAT 424
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1955 CGACGCTCAAGTCAGAGTGGGAAACCCGACAGGACTATAAGATACACAGGCGTTTCCC 2014
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 425 CGACGCTCAAGTCAGAGTGGGAAACCCGACAGGACTATAAGATACACAGGCGTTTCCC 484
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2015 CTTGGAAGCTCCCTCGTGGCTCTCTGTTCCGACCTCCGCTTACCGGATACCTGTCC 2074
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 485 CTTGGAAGCTCCCTCGTGGCTCTCTGTTCCGACCTCCGCTTACCGGATACCTGTCC 544
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2075 GCTTTCTCCCTTCGGGAAGCGTGGCGCTTCTCATAGCTCAGCTGTAGGTATCTCAGT 2134
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 545 GCTTTCTCCCTTCGGGAAGCGTGGCGCTTCTCATAGCTCAGCTGTAGGTATCTCAGT 604
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2135 TCGGTGTAGTGTGCTCCCAAGCTGGGCTGTGTGCAGCAACCCCGCTTACGCGCGAC 2194
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 605 TCGGTGTAGTGTGCTCCCAAGCTGGGCTGTGTGCAGCAACCCCGCTTACGCGCGAC 664
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2195 CGCTGGCGCTTATCCGGTAACTATCGTCTTGAGTCCAAACCCGGTAAGACACGATTCG 2254
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 665 CGCTGGCGCTTATCCGGTAACTATCGTCTTGAGTCCAAACCCGGTAAGACACGATTCG 724
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2255 CCAGTGGCAGGACCACTGTTAACAGGATTAGCAGGCGAGGTATGTAGCGGTGTCTACA 2314
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 725 CCAGTGGCAGGACCACTGTTAACAGGATTAGCAGGCGAGGTATGTAGCGGTGTCTACA 784
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2315 GAGTTCTTCAAGTGTGGCTAACTACGCTACACTAGAGGACAGTATTTGTTATCTGC 2374
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 785 GAGTTCTTCAAGTGTGGCTAACTACGCTACACTAGAGGACAGTATTTGTTATCTGC 844
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2375 GCTCTGCTGAAGCCAGTTTACCTTCGGAAAAAGAGTTG 2411
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 845 GCTCTGCTGAAG-CAGTTACCTTCGGAAAAAGAGTTG 880
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 14
LOCUS AL042026/c
DEFINITION DKF2p434E111.rl 434 (synonym: htes3) Homo sapiens cdna clone
ACCESSION DKF2p434E111.5', mRNA sequence.
VERSION AL042026
KEYWORDS AL042026.1 GI:5421372
SOURCE EST.
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 841)
AUTHORS Poustka, A., Klein, M., Mewes, H.W., Gassenhuber, J. and Wiemann, S.
TITLE EST (Poustka, et al.)
JOURNAL Unpublished (1999)
COMMENT Contact: MIPS
MIPS
```

Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@kfz-heidelberg.de;
sequenced by DKFZ (German Cancer Research Center,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No s1 sequence available.
This clone (DKFZp434E111) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

location/Qualifiers
1..841
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp434E111"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="434 (synonym: htes3)"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"

ORIGIN

Query Match 23.0%; Score 841; DB 1; Length 841;
Best Local Similarity 100.0%; Pred. No. 8.1e-236;
Matches 841; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2798 TGGCCCCAGTGTGCAATGATACCGGAGACCCAGCTCACGGCTCCAGATTATCAGC 2847
DB 841 TGGCCCCAGTGTGCAATGATACCGGAGACCCAGCTCACGGCTCCAGATTATCAGC 782
QY 2848 AATAAACACAGCCGAGGCGGAGCGGAGAGTGTCTCTGCAACTTTATCCGCCTC 2907
DB 781 AATAAACACAGCCGAGGCGGAGCGGAGAGTGTCTCTGCAACTTTATCCGCCTC 722
QY 2908 CATCCAGTCTAATTTGTCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTT 2967
DB 721 CATCCAGTCTAATTTGTCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTT 662
QY 2968 GCGCAACGTTGTGTCATGCTACAGGCATCGTGTGTCAAGCTCGTGTGGTATGCG 3027
DB 661 GCGCAACGTTGTGTCATGCTACAGGCATCGTGTGTCAAGCTCGTGTGGTATGCG 602
QY 3028 TTCATTACGTCCTGCCAAGATCAAGGCGAGTTACATGATCCCCCATGTTCTGCAA 3087
DB 601 TTCATTACGTCCTGCCAAGATCAAGGCGAGTTACATGATCCCCCATGTTCTGCAA 542
QY 3088 AAAAGCGTTAGTCTTCGGTCTCCGATCGTTGTCAAGTAAGTTGGCGCAGTGT 3147
DB 541 AAAAGCGTTAGTCTTCGGTCTCCGATCGTTGTCAAGTAAGTTGGCGCAGTGT 482
QY 3148 ATCACTCATGTTATGGCAGCACTGCATATCTTCTTACTGTCTATGCCATCGTAAGTG 3207
DB 481 ATCACTCATGTTATGGCAGCACTGCATATCTTCTTACTGTCTATGCCATCGTAAGTG 422
QY 3208 CTTTTCCTGTGACTGTGTGAGTACTCAACCAAGTCATTTCTGAGAAATAGTGTATGCGCGAC 3267
DB 421 CTTTTCCTGTGACTGTGTGAGTACTCAACCAAGTCATTTCTGAGAAATAGTGTATGCGCGAC 362
QY 3268 GAGTTGCTCTTGCCCGGCTCAATACGGGATAATACCGGCCACATAGCAGAACTTTAAA 3327
DB 361 GAGTTGCTCTTGCCCGGCTCAATACGGGATAATACCGGCCACATAGCAGAACTTTAAA 302
QY 3328 AGTGTCTATCATTTGAAACGTTCTTTCGGGGGAAACCTCTCAAGGATCTTACCGCTGTT 3387
DB 301 AGTGTCTATCATTTGAAACGTTCTTTCGGGGGAAACCTCTCAAGGATCTTACCGCTGTT 242
QY 3388 GAGATCCAGTTTCGATGTAAACCACTCGTGCACCCCACTGATCTTCAGCATCTTTTACITTT 3447
DB 241 GAGATCCAGTTTCGATGTAAACCACTCGTGCACCCCACTGATCTTCAGCATCTTTTACITTT 182
QY 3448 CACCAAGCGTTTCTGGGTGAGCAAAAAACAGGAAGGCAGAAATGCCGCAAAAAAGGGAATAG 3507

DB 181 CACCAGCGTTTCTGGTGTGAGCAAAAAACAGGAAGCAAAATGCCGCAAAAAAGGGAATAG 122
QY 3508 GCGCACAGGAATGTTGAATACTCATACTCTTCCTTTTCAATATTATTGAAGCATTTA 3567
DB 121 GCGCACAGGAATGTTGAATACTCATACTCTTCCTTTTCAATATTATTGAAGCATTTA 62
QY 3568 TCAGGGTTATTGTCTCATGAGCGGATACATATTTCAATGTATTAGAAAAATAAACAAAT 3627
DB 61 TCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAAAT 2
QY 3628 A 3628
DB 1 A 1

RESULT 15
BM438950/c
LOCUS
DEFINITION
854 bp mRNA linear EST 31-JAN-2002
IPLVr00491 Liver cDNA library Ictalurus punctatus cDNA 5', mRNA
sequence.
ACCESSION
BM438950
VERSION
BM438950.1 GI:18460672
SOURCE
EST.
ORGANISM
Ictalurus punctatus (channel catfish)
KEYWORDS
Ictalurus punctatus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes;
Ictaluridae; Ictalurus.
REFERENCE
1 (bases 1 to 854)
Feng, J., Kucuktas, H., Kocabas, A., Li, P. and Liu, Z.
Transcriptome of channel catfish (Ictalurus punctatus): initial
analysis of expressed sequence tags from the liver
Unpublished (2002)
JOURNAL
COMMENT
Contact: Liu ZJ
The Fish Molecular Genetics and Biotechnology Laboratory,
Department of Fisheries and Allied Aquacultures and Program of Cell
and Molecular Biosciences
Auburn University
203 Swingle Hall, Auburn University, Auburn, AL 36849, USA
Tel: 334 844 4054
Fax: 334 844 9208
Email: zliu@acesag.auburn.edu
Seq primer: M13 Reverse.
FEATURES
source
1..854
/organism="Ictalurus punctatus"
/mol_type="mRNA"
/db_xref="taxon:7998"
/clone_lib="Liver cDNA library"
/note="Organ: Liver; Vector: pSport1; Site_1: NotI;
Site_2: SalI"

ORIGIN
Query Match 22.5%; Score 824.6; DB 4; Length 854;
Best Local Similarity 99.3%; Pred. No. 5.6e-231;
Matches 849; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

QY 2636 TCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTCAG 2695
DB 854 TCAATCTCAAGCATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTCAG 795
QY 2696 GCACCTATCTCAGCGATCTGTCTATTTTCATCCATAGTTGCTGTGACTCCCGCTCGTG 2755
DB 794 GCACCTATCTCAGCGATCTGTCTATTTTCATCCATAGTTGCTGTGACTCCCGCTCGTG 735
QY 2756 TAGATAACTACGATACGGAGGGCTTACCAATCTGGCCCGAGTGTGCAATGATACCGCA 2815
DB 734 TAGATAACTACGATACGGAGGGCTTACCAATCTGGCCCGAGTGTGCAATGATACCGCA 675
QY 2816 GACCCACGCTCACGGCTCCAGATTATACAGCAATAAACCCAGCCGCGGAGGCCGAG 2875
DB 674 GACCCACGCTCACGGCTCCAGATTATACAGCAATAAACCCAGCCGCGGAGGCCGAG 615

QY 2876 CGCAGAAAGTGGTCTCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTCGCGGAA 2935
 DB |||||
 614 CGCAGAAAGTGGTCTCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTCGCGGAA 555
 QY 2936 GCTAGAGTAAGTAGTTGCGCAGTTAATAGTTGCGCAAGTTGTCGCAATTTGCTACAGGC 2995
 DB |||||
 554 GCTAGAGTAAGTAGTTGCGCAGTTAATAGTTGCGCAAGTTGTCGCAATTTGCTACAGGC 495
 QY 2996 ATCGTGGTGTACGCTCTGCTGTTGATGCTTTCATTCAGTCTCCGTTCCCAACGATCA 3055
 DB |||||
 494 ATCGTGGTGTACGCTCTGCTGTTGATGCTTTCATTCAGTCTCCGTTCCCAACGATCA 435
 QY 3056 AGCGGAGTTACATGATCCCGCATGTTGTCGCAAAAAAGCGTTAGTCTTCGGTCTCTCG 3115
 DB |||||
 434 AGCGGAGTTACATGATCCCGCATGTTGTCGCAAAAAAGCGTTAGTCT-CTTCGGTCTCTCG 376
 QY 3116 ATCGTGGTGTACGAGTAAGTTGCGCAGTTGATGCTTTCATTCAGTCTCCGTTCCCAACGATCA 3175
 DB |||||
 375 ATCGTGGTGTACGAGTAAGTTGCGCAGTTGATGCTTTCATTCAGTCTCCGTTCCCAACGATCA 316
 QY 3176 AATTCCTTTACTGTCATGCGCATCCGTAAGATGCTTTTCTGCTGCTGCTGCTGCTCAAC 3235
 DB |||||
 315 AATTCCTTTACTGTCATGCGCATCCGTAAGATGCTTTTCTGCTGCTGCTGCTGCTCAAC 256
 QY 3236 AAGTCATTTCTGAGAAATAGTGTATGCGCGCACCGAGTTGCTTTGCGCGCGCTCAATACGG 3295
 DB |||||
 255 AAGTCATTTCTGAGAAATAGTGTATGCGCGCACCGAGTTGCTTTGCGCGCGCTCAATACGG 196
 QY 3296 GATATACCGCGCCACATAGCAGAACTTTAAAGTGTCTCATCTGCGGAAACGTTCTTCG 3355
 DB |||||
 195 GATATACCGCGCCACATAGCAGAACTTTAAAGTGTCTCATCTGCGGAAACGTTCTTCG 136
 QY 3356 GGGCGAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCCGATGTA-CCCACTCG 3414
 DB |||||
 135 GGGCGAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCCGATGTAAGCCACTCG 76
 QY 3415 TGCACCCAACTGATCTTTCAGCATCTTTTACTTTTACCAGGTTTCTGCGTGGAGCAAAAAC 3474
 DB |||||
 75 TGCACCCAACTGATCTTTCAGCATCTTTTACTTTTACCAGGTTATCTGCGTGGAGCAAAAAC 16
 QY 3475 AGGAGGCAAAATGC 3489
 DB |||||
 15 AGGAGGCAAAATGC 1

Search completed: June 5, 2005, 22:55:16
 Job time : 10715.2 secs

This Page Blank (uspto)